



Essays on Teacher Judgements and Decisions in Ambitious Mathematics Teaching

Citation

Garcia Coppersmith, Jeannette. 2026. Essays on Teacher Judgements and Decisions in Ambitious Mathematics Teaching. Doctoral Dissertation, Harvard University Graduate School of Arts and Sciences.

Link

<https://dash.harvard.edu/handle/1/42740236>

Terms of use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material (LAA), as set forth at

<https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFIOWI2M2EwYTg>

Accessibility

<https://accessibility.huit.harvard.edu/digital-accessibility-policy>

Share Your Story

The Harvard community has made this article openly available.

Please share how this access benefits you. [Submit a story](#)

HARVARD

Kenneth C. Griffin



GRADUATE SCHOOL
OF ARTS AND SCIENCES

THESIS ACCEPTANCE CERTIFICATE

The undersigned, appointed by the

Department of Education

have examined a dissertation entitled

Essays on Teacher Judgements and Decisions in Ambitious Mathematics Teaching

presented by Jeannette Garcia Coppersmith

candidate for the degree of Doctor of Philosophy and hereby certify
that it is worthy of acceptance.

Signature Heather Hill

Typed name: Prof. Heather Hill

Signature Jon Stor

Typed name: Prof. Jon Stor

Signature Yasemin Copur-Gencturk

Typed name: Prof. Yasemin Copur-Gencturk

Signature _____

Typed name: Prof.

Signature _____

Typed name: Prof.

Date: April 2026

Essays on Teacher Judgements and Decisions in Ambitious Mathematics Teaching

A dissertation
presented by
Jeannette Garcia Coppersmith
to
The Committee on Higher Degrees in Education

in partial fulfillment of the
requirements for the degree of
Doctor of
Philosophy in
the subject of
Education

Harvard
University
Cambridge,
Massachusetts

April 2026

© 2026 Jeannette Garcia Coppersmith

All Rights Reserved.

Essays on Teacher Judgements and Decisions in Ambitious Mathematics Teaching

Abstract

Ambitious mathematics teaching involves closely attending to student sensemaking while maintaining the cognitive demand as dictated in high-level mathematics tasks. This instructional domain has been linked to the construction of deep, conceptually based subject matter knowledge. Yet classrooms serving racially minoritized students often exhibit reduced cognitive challenge and inadequate opportunities to learn. Moreover, due to its emphasis on student-generated inputs and ideas, some of which may be mathematically incorrect, teachers may be inclined to reduce the cognitive demand of high-level mathematics tasks. In this set of dissertation studies, I investigate racial-ethnic biases in teacher responsiveness to students' non-standard and innovative mathematical ideas, teachers' causal attributions for and instructional responses to student struggles on a high-level math task, and finally how teachers preserve the connectedness of the mathematics and student accountability for high-level processes and products when responding to a student misconception on a fractions task. I characterize teacher judgements and instructional preferences using vignettes to approximate practice. I first investigate teacher biases in a) expectations and judgments about student capabilities in math and b) teacher responsiveness to students' mathematical thinking. I find teachers expected greater task difficulty in both the Latinx/e and Black classroom conditions relative to the White, and they may have been more likely to support student sense-making and provide more positive, substantive affirmations to Black students relative to White students for the same mathematical solution. In the next study, I found teachers' internal—but not their external—causal attributions for a student's struggle predicted their preferences to modify the cognitive demand of high-level math tasks. Importantly, this attribution-instruction relationship was equitable across

student groups, and teachers did not show racial-ethnic biases in their causal attributions nor instructional preferences. In the final paper, I qualified teacher responses to a student misconception on a high-level fractions task in terms of the connectedness of the mathematics and student accountability for high-level processes and products. I found that teachers forged connections between mathematical ideas only part of the time, and they often took over the problem solving, diminishing the cognitive demand of the task as intended. The results of this project can inform teacher education and professional development, intervention research, and teacher evaluation to support educator capacity for delivering rigorous and just mathematics learning experiences in response to students' varied ways of thinking about mathematics.

Table of Contents

Title Page	i
Copyright	ii
Dissertation Abstract	iii
Table of Contents	v
Dedication	vi
Acknowledgements	vii
Introduction	1
Chapter 1	4
Chapter 2	61
Chapter 3	11
Conclusion	170
References	174

Dedication

I dedicate this dissertation to my late mother, Clara Angelica Garcia, and to the former students of rooms 25 and 217 who believed in the possibilities afforded by a good education. To my siblings, Nathan, and Owen, this is also for you.

Acknowledgements

The research reported here was supported in part by the National Science Foundation, the Office of the Dean at the Harvard Graduate School of Education, and the Harvard Kenneth C. Griffin Graduate School of Arts and Sciences Prize Fellowship. I also wish to thank my co-authors on the first paper: Hannah Kleen, Cynthia Pollard, and Heather C. Hill. To research associates Pawan Gupta, Arpita Tiwari, Jennifer Wu-Pope, Yu-Ping Mao, Victoria Varlack, and Cynthia Liu—thank you for your invaluable work. The opinions and ideas expressed here are those of the author and do not necessarily represent the views of the funders and associates.

Essays on Teacher Judgements and Decisions in Ambitious Mathematics Teaching

The character of math instruction outlined by National Council of Teachers of Mathematics (NCTM, 2000, 2014) and Common Core State Standards (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010) is decidedly concerned with greater cognitive activation on the part of the student. This type of instruction departs from traditional modes of mathematics teaching which emphasize didactic instruction, student memorization, and straightforward application of mathematical rules largely absent of meaning. Also termed “ambitious teaching” (Lampert et al., 2013) or “intellectually ambitious teaching and learning” (Smylie & Wenzel, 2006), this approach to math instruction centers student understandings and ideas as a resource for instruction, engages students in higher-order thinking, and supports student construction of deep subject matter knowledge in math (Smylie & Wenzel, 2006).

At its core, ambitious instruction necessitates an asset-based view of the learner—allowing a student to struggle and grapple with a math concept presupposes their inherent sensemaking abilities. There may be a pronounced expectation for student “readiness” and pre-existing abilities with consequences for ensuing instructional choices, especially for historically and racially minoritized student groups. Moreover, student struggles when learning mathematics—considered an essential aspect of the learning process and something that should be leaned into rather than avoided—may lead teachers to reduce the cognitive demand of high-level mathematics tasks during implementation in the classroom. Such judgments and decisions may lead to a vicious cycle, whereby student struggles beget remedial instruction, leading to lower mathematics proficiency and further remediation.

To better understand teacher judgements and decisions in ambitious mathematics teaching, this dissertation investigates racial-ethnic biases in teacher responsiveness to students' non-standard and innovative mathematical ideas, teachers' causal attributions for student struggles in math, and the extent to which teachers forge mathematical connections and hold the student accountable for high-level processes and products when responding to a student misconception on a high-level fractions task. The findings of this dissertation project contribute towards a better understanding of instructional mechanisms behind racial achievement disparities in mathematics education, with implications for supporting teachers in delivering rigorous and just mathematics learning experiences for all their students.

Chapter 1

Black and Latinx/e youth in urban settings too often miss out on grade-level appropriate assignments, high-quality instruction, deep engagement, and high expectations for their learning (The New Teacher Project [TNTP], 2018). In mathematics, classrooms serving racially minoritized students often exhibit reduced cognitive challenge and insufficient opportunities to learn, with a greater emphasis on procedural skill, particularly for Black and African American learners in urban settings (Boston & Wilhelm, 2017). In this study, I extend prior research on racial-ethnic biases in math teaching to include teacher responses to non-standard and even innovative student mathematical solutions. In this paper, I examine the extent to which teacher expectations, judgements and responsiveness differ between hypothetical Black, Latinx/e, and White students, with implications for racial equity in responsive math teaching.

Chapter 2

Teachers' diagnoses or explanations for student struggles relate to their practice in ambitious math teaching, especially in classrooms serving racially minoritized learners (Wilhelm et al., 2017).

However, most of the studies on teachers' causal attributions for student struggles examine them along a success-failure binary rather than the grey spaces in between, such as student stuck points during learning in progress. In this study, I examine the relationship between teachers' causal attributions for student struggles on a math task and their preferences to modify the cognitive demand of the task. I also examine the extent to which this attribution-instruction relationship is equitable across student groups. Finally, I investigate whether teachers' causal attributions and preferences to routinize, remediate, and extend a math task are influenced by student race-ethnicity, teacher background, and school context factors.

Chapter 3

Engaging students in math tasks that present a high level of cognitive demand—which enlist the student as a problem solver and forge connections between concepts, procedures, and their underlying meanings—have potential to cultivate deeper learning about fractions, yet teachers often struggle during task enactment. Adding to this, students hold enduring misconceptions about fractions, and it is unclear how teachers manage addressing them while still forging mathematical connections and enabling students to *do* mathematics as intended by high-level tasks. In this qualitative paper, I qualify teacher responses to a student misconception on a high-level fractions task in terms of the connectedness of the mathematics and student accountability for high-level processes and products. The findings have implications for research on the implementation of math tasks with high-level demands, improving teacher knowledge and practice in the face of inevitable student challenges and misconceptions, and ultimately student learning in the foundational domain of fractions.

Chapter 1

Same Idea, Shifting Standards: An Experimental Study of Racial-Ethnic Biases in Responsive Math Teaching

Abstract

Teacher expectations and judgments about student capabilities are predictive of student achievement, yet such judgments may be influenced by salient dimensions of student identity and invite biases. Moreover, responsive teaching in mathematics may also invite teacher biases due to the emphasis on student-generated inputs and ideas. While prior experimental work has demonstrated teacher biases against racially minoritized learners in the context of incorrect or partially correct student mathematical work, less is known about the potential for biases to emerge when teachers are presented with students' correct and non-standard mathematical ideas. In this pre-registered experiment, we investigate teacher biases in a) expectations and judgments about student capabilities in math and b) teacher responsiveness to students' mathematical thinking. Through a between-subjects design via an online survey, we randomly assigned $N=312$ teachers to a simulated classroom composed of predominantly Black, Latinx/e, or White students and prompted them to respond to six vignettes featuring correct and non-standard student solutions to mathematics problems. Our findings show teachers expected greater task difficulty in both the Latinx/e and Black classroom conditions relative to the White. We did not find significant differences by condition in other dimensions, though we found trend-level evidence to suggest teachers may be more likely to support student sense-making and provide more positive, substantive affirmations to Black students relative to White students for the same mathematical solution. Our findings have implications for research and teacher training in reform-oriented mathematics instruction.

Same Idea, Shifting Standards: An Experimental Study of Racial-Ethnic Biases in Responsive Math Teaching

“Teachers, like all of us, use the dimensions of class, race, sex, ethnicity to bring order to their perception of the classroom environment” (Lightfoot, 1978, pp. 85-86)

The United States has one of the widest math performance distributions in the world (National Center for Education Statistics, 2019). European American or White¹ and Asian American students consistently demonstrate stronger math achievement compared to Latinx/e², Native American, and Black³ students, contributing to magnified disparities in later educational outcomes, lifetime earnings, and overall access to opportunity in American society (Gutiérrez, 2012; Jencks & Phillips, 2011; Reardon et al., 2015). Scholars have called for a critical examination of “opportunity gaps” for student learning in math, gaps which may help to explain differences in student performance (Flores, 2007; Ladson-Billings, 2006; Martin, 2009; Milner, 2010). Such gaps in opportunity include differences in overall teacher quality, curricula, school funding, as well as educator expectations and mindsets about students of color that delimit access to academically rigorous learning experiences (Irvine, 2010; Milner, 2012).

To better understand opportunity gaps, scholars have called for a closer examination of racial-ethnic⁴ biases in teachers’ instructional decisions (Warikoo et al., 2016). Indeed, disparities

¹ In line with prior scholarship by Ewing (2020) and others, we capitalize "White" in order to underscore "the specificity and significance of Whiteness—the things that it is, the things that it does."

² We use the gender-neutral and inclusive pan-ethnic identifier “Latinx/e” while understanding that not all members of this group may self-identify with this term (Trejos, 2024).

³ We use the term “Black” in order to recognize this racial-ethnic identity in a transnational manner, inclusive of individuals with roots in the Caribbean and Latin America (Laws, 2020).

⁴ Our use of the term "racial-ethnic" aligns with prior scholarship on the interrelationship between racial and ethnic identities. We take race to define a "socially constructed lay theory" (p. 4) representing perceived phenotypic differences between groups of people. Similarly, ethnicity represents national origin and the cultural traditions of social groups. Race and ethnicity are socially constructed and highly correlated both in terms of perception and in terms of sociocultural identity. Further, while terms like "Latinx" represent an ethnic social categorization, terms like "Black" often represent both racial and ethnic social categories. As such, we take the term “racial-ethnic” to capture the overlapping meanings of these constructs (Cross & Cross, 2008).

between Black and White student achievement in the U.S.—an important measure of education inequality—correlate with teachers’ implicit racial biases at the county level, suggesting a relationship between teacher bias, instructional mechanisms, and student achievement (Chin et al., 2020).

Findings from audit, experimental, and naturalistic studies have demonstrated racial-ethnic biases favoring dominant student groups, with consistent findings across various instructional dimensions identifying biases in teacher’s general academic expectations (Tenenbaum & Ruck, 2007; van den Bergh et al., 2010) and associated student achievement (Peterson et al., 2016). Biases against racially minoritized learners have also been found in teachers’ assessments of student ability for ambiguous solutions in math (Copur-Gencturk et al., 2019), grading (Malouff & Thorsteinsson, 2016; Quinn, 2020b), feedback and praise for writing (Harber et al., 2012), general classroom discourse patterns, and disciplinary and educational service referrals (Tenenbaum & Ruck, 2007).

We extend this work in three ways. First, to our knowledge, no experimental studies have investigated biases in teacher expectations of task difficulty—teachers’ prospective judgments about how hard a task will be for students—in the context of correct, non-standard, and innovative student math solutions. Critically, teacher expectations not only predict student achievement, they also vary between students of different race/ethnicities, with lower expectations for students belonging to historically racially minoritized groups (Peterson et al., 2016; Pit-ten Cate & Glock, 2023; Rubie-Davies et al., 2006; Weinstein, 2002). We argue that assigning tasks to students is a concrete mechanism by which expectations can introduce bias into students’ opportunities to learn and investigate those difficulty expectations here. Indeed, prior research has found racial bias against Black students in teacher task selection (Pollard, 2022), making difficulty estimations a fruitful line of study.

Second, few studies have investigated whether biases arise in teachers' in-the-moment decisions about instruction, despite teachers having wide discretion about such decisions (Ball, 2018). In mathematics, the focus of this study, teacher biases may be particularly relevant to responsive teaching, in which the teacher leverages the substance of students' mathematical thinking to guide instruction (Carpenter et al., 1989; Franke et al., 2009; Jacobs et al., 2010). Responsive teaching in mathematics includes instructional moves such as eliciting student thinking, using that thinking in subsequent instruction, and affirming the substance of students' mathematical ideas (Kang, 2022; Richards & Robertson et al., 2015). In such classrooms, student sense-making about mathematics is "put into play" as a resource for learning, and teachers respond by affirming and building on those ideas. Attending and responding to the substance of students' mathematical ideas may reduce the potential for biases, as this provides the teacher with more information and may leave less room for stereotype-based inferences (Bertrand & Duflo, 2016). However, as Ball (2018) notes, the high volume of teacher decisions when responding to students during math instruction—and the wide latitude for discretion—may create more opportunities for teacher biases to emerge.

Third, the majority of scholarship on racial-ethnic biases in mathematics education has focused on Black students and White students, and we extend work in this area to include Latinx/e students. Latinx/e students comprise the fastest growing and second largest racial-ethnic group in U.S. public schools, making it crucial and relevant to include them in any examination of biases against racially minoritized learners (National Center for Education Statistics, 2024).

We thus investigate the relationship between student race and teacher expectations and responsiveness. Specifically, we randomly assigned teachers to a simulated classroom environment composed of predominantly Black, Latinx/e, or White students and then prompted them to verbally respond to a hypothetical target student's mathematical solution. We also prompted teachers to rate

the quality of the hypothetical student's mathematical thinking and the level of difficulty of the problem for the typical 4th grader in the U.S. Thus, we ask: What is the extent to which teachers demonstrate biases amongst Black, Latinx/e and White students in their...

- 1) ... expectations and judgments about student math capabilities, including a) teacher judgments about the quality of student mathematical thinking and b) teacher expectations of task difficulty for typical students?
- 2) ... instructional moves that are responsive to students' mathematical solutions, including elicitation and use of student thinking, substantive affirmations to students, and student-centered instructional moves?

We argue that this study combines the clear causality that arises from an experimental design with more authentic depictions of classroom life and teacher decision-making (Herbst et al., 2011; Charalambous, 2020). Because responsive instruction and high expectations are linked to improved learning outcomes for students (e.g. Rubie-Davies et al., 2006; Smylie & Wenzel, 2006), information about the extent to which student racial-ethnic groups may experience such opportunities to learn differentially in these high-leverage instructional domains will contribute to the field's understanding of how racial achievement disparities come to be at the classroom level. More importantly, a concrete understanding of the instructional mechanisms behind racial inequality has the potential to inform teacher education and professional development with specific areas to intervene on, for example, through teacher coaching and rehearsals (Mancenido et al., 2025).

Background

We begin by outlining our theoretical framework based on implicit social cognition. We follow with a discussion of the prior literature on racial-ethnic biases in teacher expectations and instruction with particular attention to domain-general biases and those specific to the mathematics

instructional context in which teachers respond to incorrect, partially correct, and correct student mathematical work. We follow with an overview of responsive teaching in mathematics and its relationship to student learning, arguing that this type of teaching may simultaneously instigate and inhibit student opportunities to learn according to how these dimensions of instruction interact with teacher biases. We articulate the gap we aim to address before arriving at the present study.

Theoretical Framework for Racial-Ethnic Biases in the Classroom

An examination of racial-ethnic biases in the classroom through the lens of social psychology may help to explain disparities in student opportunities to learn (Warikoo et al., 2016). Specifically, implicit social cognition theory, which refers to the subconscious influence of one's past experiences on their perceptions, attitudes, and behaviors (Greenwald & Banaji, 1995), provides a theoretical framework for understanding the ways in which stereotypes may shape teachers' cognition and instructional decisions. Implicit biases are the "stereotyped beliefs or prejudices that occur without conscious control or even necessarily the awareness of the individual who holds them" (Yogeeswaran et al., 2018, p. 70). Explicit biases, on the contrary, are conscious and need controlled reflection to be activated (Fazio, 1990). Especially when it comes to implicit stereotypes, people infer information mostly on the basis of apparent social categories.

Biases in behavior and judgments trace back to stereotypes or cultural beliefs about the behavior and traits of social groups (Eagly & Chaiken, 1993; Hilton & von Hippel, 1996; Tajfel, 1970). A core function of stereotypes is group categorization, which is used to save time and cognitive resources when processing information and making sense of one's complex social world (Schneider, 2004). As a result of stereotyped thinking, individuals in a social group tend to be perceived in accordance with the stereotypes about that social group. Such associations are overgeneralized and usually develop through socio-cultural factors, socialization processes, media

consumption, encounters with cultural narratives, and actual experiences with social groups (Dovidio et al., 2001). According to this theory, salient identity features such as race-ethnicity come into play when an individual processes social information, activating stereotypes. The effect of stereotypes on individual cognitive processing is biases in behavior and judgments, which are not based on the actual competencies and characteristics of a person, but rather on their group membership. Moreover, in the absence of information, actors may rely more on group membership when making assessments about an individual (Bertrand & Duflo, 2016). The influence of stereotyped associations is stronger when one is given greater latitude for interpretation: when there is a lack of information, group association works to “fill in the blanks” based on the information readily available from stereotypes (Bertrand & Duflo, 2016; Dovidio, 2001). Thus, stereotypes can become more influential in low-information and ambiguous situations, such as when confronting non-standard, unconventional, or incomplete student ideas in the classroom.

In education, dominant social narratives about Black and Latinx/e students have the potential to inform stereotyped perceptions of minoritized learners, with consequences for instruction. Racial-ethnic minority groups have been harmfully mischaracterized as having low intelligence or being less hard working (Bonefeld et al., 2020; Fiske et al., 2007). Stereotypes about minoritized learners may indeed shape individual’s perceptions; in a randomized experiment, Quinn (2020a) found that viewers who watched a TV news story about racial achievement gaps expressed more exaggerated stereotypes of Black Americans as lacking education and showed increased implicit stereotyping of Black students as less competent compared to their White counterparts. Further, negative stereotypes have the potential to shape instructional decision making. In their study of pre-service math teachers’ engagement in a mixed-reality simulation, Bondurant and Reinholz (2023) found that the (White) actors playing the role of students reinforced negative racial and gender stereotypes about students, and pre-service teachers in turn provided lower-quality

engagement opportunities (e.g. “why” questions) to negatively stereotyped students. Common stereotypes in education concerning minoritized learners, particularly in math, are enduring and have the potential to shape teacher perspectives of their students and decisions in the classroom (Martin, 2009).

Paradoxically, stereotypes that lead teachers to hold lower expectations may result in more positive judgments about specific individuals’ work. This occurs, as described by the shifting standards theory, when stereotypes about social groups inform standards for evaluating members of that group, and those standards shift and differ across groups (Biernat et al. 1991; Biernat, 1995). Marshall’s (2020) discussion of the use of terms like “well-spoken” or “articulate” in reference to Black people simply communicating effectively helps to illustrate the shifting standards theory. In education, the shifting standards theory is evident in qualitative work showing teachers assume a background of hardship for their Black and Brown learners and thus “relax their standards for these students because they feel sorry for them” (Benson & Fiarman, 2020, p. 83). This phenomenon implies a low standard for members of racially minoritized groups and result in more positive judgments about the same skill or ability as compared to those for White individuals.

Biases in Teachers’ Academic Expectations and Judgements About Student Abilities

Teacher expectations are the cognitive, “inferential judgments that teachers make about probable future achievement and behavior of a student” (Brophy & Good, 1974, p. 129). In mathematics, expectations and judgments about student abilities can involve teacher expectations of the difficulty of a problem or task for a student and judgments about the quality and sophistication of students’ mathematical thinking. Informed by implicit stereotyped perceptions, teacher expectations differ across student groups, as they are influenced by characteristics such as the student’s socioeconomic status, gender, (dis)ability, language, and race-ethnicity—the focus of this

study (Rubie-Davies et al., 2012). Prior research provides evidence on this point: a meta-analysis of 32 studies concluded teacher expectations are higher for White students compared to Black ($d=.25$) and Latinx/e ($d=.46$) students (Tenenbaum & Ruck, 2007). Non-Black teachers have also been found to hold significantly lower expectations for Black students compared to Black teachers (Gershenson et al., 2016). Further, an experimental audit study by Anderson-Clark et al. (2008) found that teacher ratings of students' motivation and achievement-related behaviors were significantly lower for a Black-associated name compared to a White-associated name, but no main effect was found from explicitly-stated descriptions of student race. These studies highlight the unique role of implicit racial biases over explicit ones in influencing teacher expectations for and judgments about student academic abilities.

Such differential expectations can be consequential for student achievement. Classic studies from the 1960's and '70's describe the self-fulfilling, prophetic nature of teacher expectations in elementary school—both for leading to higher student achievement when teachers held positive expectations about students' intellectual abilities (Rosenthal & Jacobson, 1968) or to lower student achievement when teachers held negative expectations based on student social class (Rist, 1970). When it comes to race-ethnicity, teachers' implicit biases have been shown to predict racial achievement gaps between Dutch-origin and Turkish immigrant students, and this relationship is mediated by teacher expectations (van den Bergh et al., 2010). Peterson and colleagues (2016) found implicit biases in favor of ethnic-majority student groups in New Zealand likewise predicted student achievement in reading and math, and that students in classrooms of teachers who explicitly held higher expectations for their learning showed greater improvements in reading. While some scholars have argued that differential teacher expectations by student race merely reflect racialized achievement disparities (Jussim & Harber, 2005), Peterson and colleagues' models (2016) identified the effect of bias while also controlling for prior student achievement.

In descriptive work in the domain of mathematics specifically, we find mixed evidence of biases in teacher expectations and judgments against racially minoritized learners. Qualitative and ethnographic studies have documented low teacher expectations for Black and Latinx/e students from the perspectives of the students themselves, who have shared feelings of not being challenged and needing to prove themselves in math class (Davis, 2014; Gross, 1993; Leyva 2021). These qualitative findings illustrate the low expectations too often experienced by racially minoritized students in mathematics. However, in a descriptive study using a nationally representative sample, Riegle-Crumb & Humphries (2012) examined biases in teachers' assessments of math course difficulty for their students, finding that after accounting for student performance, differences in teacher assessments of course difficulty against Black and Hispanic students disappeared, and that biases against White girls (relative to White boys) remained. Lacking in this study is a consideration of how course-level and achievement may themselves relate to biases, raising the need for experimental work.

In the few experimental studies on teacher expectations and judgments about student abilities in the domain of math, findings are mixed but point to some biases against racially minoritized learners when responding to incorrect and partially correct student work—an area of ambiguity that may instigate stereotyped associations. Using a fractions task involving a common error when adding unlike fractions, Pollard (2022) found teachers were more likely to attribute errors to the student, rather than to the student's prior instruction, for Black students as compared to White students. Additionally, Pollard (2022) found this effect was larger when the teacher held stronger pro-White implicit bias (as measured by the Teachers' Implicit Association of Academic Achievement Task (TIAAAT) adapted from Peterson et al. (2016)). Examining interactions between student gender and race-ethnicity, Copur-Gencturk and colleagues (2019) found that while teachers were unbiased when rating the correctness of student solutions, they demonstrated racial-ethnic

biases in their ratings of student math abilities for partially correct solutions, an area of ambiguity. Teacher ratings were lowest for Black and Latinx/e or Hispanic girls compared to boys and White students, underscoring the additive effects of marginalized identities for inviting biases. However, no evidence of racial biases was found on average in a study that assigned race-specific names to partially correct student math solutions (Copur-Gencturk et al., 2023) or in a similar experiment conducted in the virtual teaching space, which also found no evidence of correctness or ability bias from teachers when viewing incorrect or partially correct work (Copur-Gencturk et al., 2022). These experimental findings demonstrate the inconsistencies of biases in teachers' subjective judgments about students' mathematical abilities, particularly at the intersections between marginalized identities of student race and gender when viewing partially correct or ambiguous student work in math.

There is less experimental work examining racial biases in teacher judgments and expectations when responding to correct student mathematical work. Battey and colleagues (2021) examined pre-service teachers' racial attitudes and perceptions of student thinking when viewing two different but comparable videos featuring Black or White students correctly explaining equivalent fractions. They found that pre-service teachers rated the quality of students' mathematical thinking more favorably and used more positive language to describe mathematical strategies and thinking for White students as compared to Black students, and that these biases related to their explicit (but not implicit) racial attitudes and interracial contact with Black communities. Because this particular study did not use random assignment, it is not clear whether student race-ethnicity or sociocultural factors related to students' ways of knowing and speaking about math (Nasir et al., 2008) triggered teacher biases. These findings underscore the need for experimental work to disentangle predictors of teacher bias in this area, more work investigating biases against Latinx/e

students in addition to Black students, and more studies of biases in the context of mathematically correct and non-standard student solutions.

Biases in Teachers' Instruction

Biases have also been shown in various dimensions of teachers' instruction. In their meta-analysis of both experimental and non-experimental studies on teacher speech patterns, Tenenbaum and Ruck (2007) found racial biases in teachers' valence of speech and referrals, with more positive speech (e.g. encouragement) and more neutral speech (e.g. questions) for European American students as compared to Black and Latinx/e students. Teachers were found to direct negative speech equally across student groups. While the associations were small, they were statistically significant within the meta-analytic framework, in line with previous research on the small-to-moderate effects of implicit biases. However moderate the effect sizes may be, the consequences of racial-ethnic biases have the potential to be cumulative and harmful (Greenwald et al., 2014) over the course of a student's entire K-12 schooling history, leaving them underprepared for higher education.

Scholarship has also uncovered racial biases in teachers' subjective feedback on student writing. In an experimental audit study of student essays, Harber et al. (2012) found that White teachers demonstrated positive feedback bias toward Black and Latinx/e students, providing less critical feedback than they did for White students, but only on the more subjective components of writing like content rather than mechanics. In line with the shifting standards theory described above, the authors of this study hypothesize that teachers may have provided more positive affirmations on Black and Latinx/e students' writing because they had lower initial expectations for their writing skills. While prior research shows teachers provide biased ability ratings towards Black

and Hispanic students' work (Copur-Gencturk et al., 2019), there are no studies to date, to our knowledge, that examine biases in teachers' affirmations of student thinking in mathematics.

Experimental research on racial biases in math teaching has begun to emerge, with—similar to the above—most of the findings identifying teacher biases against Black relative to White students in the context of instructional responses to incorrect or partially correct student mathematical work. In a randomized experiment with 90 general education mathematics teachers in grades 4-8, Pollard (2022) focused on the data-use cycle—where teachers evaluate student work, diagnose errors, provide feedback, and make curricular and pedagogical choices for instructional remediation (Datnow & Hubbard, 2015). Using a hypothetical roster containing student demographic information, teachers were randomized into a predominantly White or Black classroom, viewed a work sample from a hypothetical Black or White boy containing a mathematical error adding unlike fractions, provided written feedback to the student, and finally selected ensuing curricular and pedagogical choices. Pollard (2022) found that teachers were more likely to select lower-quality math curricular materials and remedial instruction for Black students compared to White students and that teachers were more likely to endorse student-centered instructional moves for Black students compared to White students. However, student race-ethnicity showed no relation to the evaluation or feedback teachers provided. Further, Copur-Gencturk and colleagues (2022) did not find any racial-ethnic biases in teacher recommendations for gifted education based on correct or mostly correct student mathematical work but did find that teachers were more likely to refer Black students for special education in the context of incorrect and partially correct student solutions. These studies underscore the role of racial-ethnic biases in teachers' curricular and pedagogical choices and their referrals in the context of incorrect mathematical ideas. We build on this work by similarly looking at biases but include Latinx/e in addition to Black students and shift

the focus from the choice of tasks to teacher responsiveness to mathematically correct and non-standard student ideas.

Findings demonstrating teacher biases in response to student work containing some math errors align with prior results from audit experiments in psychology in which biases have been found to inform behavior and judgments when respondents are given greater latitude for interpretation—in other words, discretion. Indeed, in the absence of information and in ambiguous situations that require subjectivity, student group membership may become more salient, causing implicit biases to play a more pronounced role in informing action (Dovidio, 2001). In responsive teaching—where teachers must interpret student thinking and articulate a reasoned reaction in the moment (Jacobs et al., 2010)—implicit biases may be more likely to affect teachers’ cognitive processing. Further, models of responsive teaching (Alexander, 2008; Lampert, 2001; see below for more detail) encourage students to express their mathematical ideas and strategies, often resulting in correct but non-standard student solutions, another area of ambiguity requiring teacher discretion. Yet we know of no studies to date that examine racialized differences in teacher responsiveness in this crucial instructional domain.

Responsive Teaching in Mathematics

Responsive teaching—sometimes called dialogic teaching (Alexander, 2008)—leverages student thinking as an instructional resource to build on students’ mathematical knowledge. In line with reforms outlined by National Council of Teachers of Mathematics (NCTM, 2000, 2014) and Common Core State Standards Initiative (2010), this approach to instruction centers students’ mathematical ideas by inviting them to explain and justify their thinking processes and attend to each other’s reasoning. When teaching responsively, teachers must attend to, interpret, and respond to the mathematical details in students’ strategies (Jacobs et al., 2010). Leveraging student

sensemaking in this way is associated with improved learning outcomes in math, as it supports higher-order thinking and the construction of conceptually based subject matter knowledge (Carpenter et al., 1989; Cohen & Hill, 2000; Schoenfeld et al., 2023; Smylie & Wenzel, 2006).

Effective responsive teaching in mathematics necessitates that teachers make a variety of instructional decisions as they work with student mathematical thinking in the moment (Boston, 2012; Stein et al., 2008). These decisions may include (but are not limited to) eliciting, using, and affirming that thinking as well as supporting student sensemaking about the math content. *Elicitation* refers to the practice of surfacing specific students' mathematical ideas and can involve prompting a student to explain their strategies or other ways of "getting student thinking on the table" (Shaughnessy & Boerst, 2018). *Use of student thinking* in mathematics refers to the teacher keeping a student's idea in play as an object of discourse and includes the teacher or another student revoicing, emphasizing, restating, and probing a student's idea or solution process (Herbel-Eisenmann et al., 2013; Jacobs & Empson, 2016; Lampert, 2001; O'Connor & Michaels, 1996). *Affirming* student thinking includes teacher praise and "affirmation of something a student did well" (Brophy, 1981; Sun & Ruef, 2023). Finally, *supporting student sensemaking* refers to instructional moves in mathematics that create opportunities for students to have ownership over, grapple with, and make sense of the content (Schoenfeld et al., 2023). As opposed to the teacher driving the mathematical thinking, for example, by prescribing a specific mathematical strategy or step-by-step instructions, teachers can support student sensemaking by creating opportunities for learners to bring their existing knowledge to bear and make sense of mathematics. This could look like engaging students in a think-pair-share or having them compare their mathematical strategies.

As evidenced by the variety of instructional moves teachers can choose to deploy in responding to and assessing student thinking in the moment, responsive teaching opens the door to

the broad realm of student ideas and ways of thinking about mathematics, introducing greater possibilities for instructional content and consequently a greater degree of uncertainty. Greater attention to student thinking is necessitated, as teachers must attend to students' sometimes unconventional intellectual contributions as well as to the established standards of the discipline (Lampert, 1990). On one hand, surfacing the array of potential student ideas in the math classroom can create new opportunities to reveal student thinking that would otherwise go unacknowledged, making learning experiences more equitable. On the other hand, contributions from minoritized learners may be undervalued or go unnoticed in the interactional space of the math classroom, thereby reproducing racial hierarchies (Battey et al, 2021; Gutiérrez, 2013; Louie, 2018; Martin, 2009; Nasir et al., 2008). For example, in response to a problem asking, "*How many tens does 743 have?*", a non-standard student approach might be "*seventy-four tens*," which a teacher might interpret as incorrect, technically correct but not relevant to the question, or mathematically novel and innovative. Depending on how the teacher perceives this approach and whether the student comes from a racially minoritized background, the teacher may judge this student to be mathematically "exceptional," performing at the expected level, or in need of a review of place value. Thus, responsive teaching may instigate biases if a non-standard response is seen as a lack of knowledge or as a sign of mathematical "exceptionalism" for racially minoritized students.

Emerging research shows mixed findings about the potential for reform-oriented mathematics teaching to rectify social inequities. In some undergraduate math classrooms focused on inquiry-based discussion, women's participation was found to be significantly lower than men's, which later predicted gender gaps in mathematics performance in favor of men (Reinholz et al., 2022). However, Laursen et al. (2014) found that inquiry-based learning in college mathematics eliminated gender gaps found in traditionally taught courses. Moreover, in a descriptive study by Reinholz and Wilhelm (2024) examining student talk turns in a racially diverse urban district that

had implemented mathematics instructional reforms, Black students were found to have high levels of participation, indicating what the authors describe as a sign of improved equity in the district. However, they also found Latinx/e students showed significantly lower levels of participation compared to Black and White students, raising questions about the potential for reform-oriented mathematics teaching to rectify and/or exacerbate racial disparities in mathematics teaching and learning.

The Present Study

Experimental findings from the literature on racial-ethnic biases in the domain of mathematics have primarily focused on incorrect or partially correct student work and have often demonstrated lower expectations and perceptions of student ability against Black students, work we replicate here and extend in three important ways. First, we aim to reproduce prior experimental findings regarding racial biases in teacher expectations and judgments about student ability and extend them to contexts featuring correct, non-standard, and innovative student math solutions. Second, we extend prior studies about biases in instruction to key elements of responsive teaching. Finally, we investigate biases against Latinx/e students in addition to Black students. We do this work using a pre-registered experimental design, strengthening the study's internal validity.

Methods

We measured teacher judgments about the quality of students' mathematical thinking, teacher expectation of task difficulty, and teacher responsiveness to student solutions in the context of six animated teaching simulations, or ATSs (Herbst et al., 2011; Charalambous, 2020), here called vignettes. We then ran a series of regressions on these outcomes of interest to understand how teachers' responses to and thinking about student work in the vignettes might systematically vary depending on student race-ethnicity. Although we may sacrifice the face validity associated with

classroom observation data with our use of vignettes, this approach allows us to make causal inferences about the role student race-ethnicity may play in teachers' expectations of students and their subsequent instructional decisions. The use of vignettes also allowed us to reach a large sample of teachers, necessary for conducting this experimental work. We emphasize that claims about racial bias tend to carry a greater "burden of proof" in the social and psychological sciences (e.g. Roberts, 2022). As race tends to correlate with factors like socioeconomic status and school resources, including teacher quality (e.g. Clotfelter et al., 2023), we elected for an experimental design and varied student race-ethnicity. Our experimental methodological approach enables us to make inferences about the specific role racial-ethnic biases play in mathematics instruction. We describe our methods in more detail next.

Sample

We elected to work with K-8 teachers because these years are critical, as they set up the foundation for algebraic thinking and higher-level mathematics, and because this large population of teachers enabled us to achieve a sufficient sample. We recruited participants using a Qualtrics teacher panel in the U.S. from June through August of 2022. Our screening criteria selected for K-8 teachers who had taught mathematics within the last three years. Qualtrics returned a data set of 315 teachers, from which we omitted three teachers who failed to authentically engage in the survey (e.g. had blank, incomplete, or irrelevant responses). While our study focuses on K-8 teachers, we retained the 19 teachers who reported they taught high school in 2022 because our screening criteria was inclusive of any teacher who had taught K-8 mathematics in the prior three years. Teachers with secondary mathematics credentials could have switched from teaching middle grades to high school math in that three-year window. Our survey included attention checks throughout to check if participants were paying careful attention (e.g. "How many students have their hands raised?"), with

over 99% of teachers passing at least one. Because an inspection of responses revealed that individuals who failed an attention check did offer appropriate and relevant verbal responses to the vignettes, we did not remove their responses. Our final sample of $N=312$ teachers is shown in Table 1.1. In line with the teacher population in the U.S., the majority of our sample identified as White women. Our sample, however, is weighted toward more experienced educators than is typical in the U.S. (National Center for Education Statistics, 2023).

Table 1.1 Sample Characteristics for Full Sample and Across Conditions

	Full Sample $N=312$	Condition A $n=109$	Condition B $n=106$	Condition C $n=97$	Pearson Chi-square or ANOVA
	N (%)	n (%)	n (%)	n (%)	
Gender					
Woman	272 (87)	95 (87)	88 (83)	89 (92)	$\chi^2(2) = 3.46, p = 0.18$
Man	39 (13)	14 (13)	18 (17)	7 (7)	$\chi^2(2) = 4.43, p = 0.11$
Nonbinary	1 (<1)	0 (0)	0 (0)	1 (1)	$\chi^2(2) = 2.22, p = 0.31$
Teaching experience					$\chi^2(12) = 10.52, p = 0.57$
<1 year	3 (1)	1 (1)	2 (2)	0 (0)	
1-2 years	18 (6)	7 (6)	5 (5)	6 (6)	
3-5 years	34 (11)	14 (13)	6 (6)	14 (14)	
6-10 years	66 (21)	25 (23)	22 (21)	19 (20)	
11-15 years	46 (15)	18 (17)	18 (17)	10 (10)	
16-20 years	41 (13)	13 (12)	13 (12)	15 (15)	
20+ years	105 (34)	31 (28)	40 (38)	33 (34)	
Lowest Grade taught					$\chi^2(6) = 4.13, p = 0.66$
PreK-2	82 (26)	24 (22)	34 (32)	24 (25)	
3-5	139 (45)	51 (47)	45 (42)	43 (44)	
6-8	72 (23)	28 (26)	22 (21)	22 (23)	
9-12	19 (6)	6 (6)	5 (5)	8 (8)	

Table 1.1 Continued

	Full Sample	Condition A	Condition B	Condition C	Pearson Chi-square or ANOVA
	N=312	n=109	n=106	n=97	
	N (%)	n(%)	n(%)	n(%)	
Teacher race-ethnicity					
Black	22 (7)	3 (3)	14 (13)	5 (5)	$\chi^2(2) = 9.74, p = 0.01$
Latinx/e	16 (5)	7 (6)	4 (4)	5 (5)	$\chi^2(2) = 0.78, p = 0.69$
Native Am.	3 (1)	0 (0)	1 (1)	2 (2)	$\chi^2(2) = 2.29, p = 0.21$
AAPI	11 (4)	2 (2)	5 (5)	4 (4)	$\chi^2(2) = 1.46, p = 0.51$
White	239 (77)	88 (81)	73 (69)	78 (80)	$\chi^2(2) = 5.36, p = 0.07$
Multiracial	21 (7)	9 (8)	9 (8)	3 (3)	$\chi^2(2) = 2.97, p = 0.20$
School Demographics	mean				
Urbanicity	0.29	0.31	0.30	0.27	$H(2) = 2.22, p = 0.33$
Percent White	0.55	0.55	0.53	0.57	$F(2, 305) = 1.16, p = 0.32$
SES average	0.28	0.32	0.20	0.34	$F(2, 306) = 2.48, p = 0.09$
Avg. Math Ach	0.03	0.02	-0.005	0.06	$F(2, 305) = 2.62, p = 0.07$

Note: Experimental conditions A, B, and C consisted of a hypothetical classroom predominantly composed of Black, White, or Latinx/e students, respectively. AAPI stands for Asian American and Pacific Islander. Column percentages are shown in parentheses. Means are provided for school demographic variables. We ran one-way ANOVA tests to compare means across conditions for school demographic variables (for urbanicity, which showed a non-normal distribution, we used a Kruskal-Wallis test). For our categorical variables, we used Chi-Square tests and report the Fisher's exact test results when cells had few (<5) frequencies.

Study Design

We employed a between-subjects design in which participants were each randomly assigned to view six vignettes from a hypothetical elementary classroom composed of predominantly Black students (condition A, $n=109$ teachers), White students (condition B, $n=106$ teachers), or Latinx/e students (condition C, $n=97$ teachers). Randomly assigning teachers to experience vignettes in only

one condition rather than across conditions (a between- rather than within- subject design) diminished the potential for social desirability bias or participants becoming aware that the study was about race (Hofmann et al., 2005). Vignettes were identical save for the race-ethnicity of the students. To test for potential differences in teacher characteristics across our three conditions, we ran a series of Chi-Square and ANOVA tests. We found that by chance, Black teachers were over-represented in the condition predominantly composed of White students ($\chi^2(2) = 9.74, p = 0.01$), and White teachers may have been slightly over-represented in the condition predominantly composed of Black students ($\chi^2(2) = 5.36, p = 0.07$). We also found that teachers in the White classroom condition may have worked in metro locales with a moderately lower socioeconomic status ($F(2, 306) = 2.48, p = 0.09$) and lower aggregate mathematics achievement ($F(2, 305) = 2.62, p = 0.07$) compared to teachers in the other two conditions. We discuss these imbalances in our limitations section but note that our analytic models account for teacher race-ethnicity and school context factors as covariates, meaning we interpret differences in our dependent variables across conditions above and beyond these teacher- and school-level factors. We found no statistically significant differences between groups of teachers across the three conditions for our other demographic and contextual variables.

We composed vignettes of between 5 to 11 slides, or frames. To inform the development of vignettes, we conducted cognitive interviews with elementary teachers and fine-tuned their content and format before piloting them on the online survey platform (Hill et al., in press). Each vignette showed study participants an elementary math problem introduced by a ‘teacher;’ student work time on the problem; and a “target” student solution to the problem which provided either their answer, their thinking, or both. All target student responses could be interpreted as correct, and five of the six featured non-standard—and sometimes innovative and sophisticated—student mathematical reasoning. As the relevant target student utterances in Table 1.2 show, all responses invited further

probing or discussion of the student's thinking. The final frame of the vignette first asked study participants to verbalize what they would say to the class next into their device's microphone, and then to, in a second step, verbalize the instructional step they would take after that initial comment.

Table 1.2. Item Descriptions and Target Student Responses in the Six Vignettes

Item	Mathematical problem or prompt	Instructional Scenario	Relevant target student utterances
1	Represent $\frac{2}{3}$ on a number line	Students work independently. Teacher brings class together and asks two students to share their answers.	S1: First I divided the part of the number line between 0 and 1 into 3 equal segments. Then, I shaded two of them. So $\frac{2}{3}$ is right here. S2: I did the same thing. I just shaded in two different segments. (Shows shading starting at $\frac{1}{3}$ and ending at 1).
2	T has 14 papers. J has 9 papers. How many fewer papers does J have than T?	Students work independently. Teacher asks a student to share thinking. Student replies using "counting up" method.	S: I used my fingers and counted up from 9 to 14. That's 10, 11, 12, 13, 14. That's 5 counts, so J has 5 papers fewer than T.
3	What fraction of the rectangle is shaded? (Rectangle shows $\frac{2}{5}$ shaded)	Students think independently and then turn and talk about problem. Teacher elicits a response from a student. Student replies with a numeric answer and a solution.	S: Two fifths. There are five parts and two of them are shaded.
4	$12 \times 5 =$	Teacher elicits student response. Student replies with a non-standard method.	S: I did it in my head. First I multiplied 5 times 10 and got 50. Then I multiplied 5 times 2 and got 10. Then I added 50 and 10 and got 60.

Table 1.2. Continued

Item	Mathematical problem or prompt	Instructional Scenario	Relevant target student utterances
5	$19 + 6 =$	Teacher asks for a student to come to the board to solve problem. Student volunteers and solves the problem using a non-standard method.	<i>S</i> : I took away a one from the 6 and added it to the 19. Then it was easy, 5 plus 20 is 25.
6	How many tens does 743 have?	Teacher first asks how many ones 743 has. Student answers 3, and teacher affirms answer as correct. Teacher then asks how many tens 743 has. Teacher calls on two students.	<i>S1</i> : I think it has four tens. <i>S2</i> : I think it has seventy-four tens.

Note: To view vignettes, see https://osf.io/qz5g4/overview?view_only=fecd5b01074146b98765af1f4444f617

To operationalize race-ethnicity, four of the six students present in each vignette signaled either Latinx/e, Black, or White students in terms of physical attributes like skin tone and hair features. Further, we assigned a name to the student speaking that was commonly associated with that particular group, but not socioeconomic status, to further signal their racial-ethnic identity (Gaddis, 2017a, b). As is typical for U.S. classrooms, the teacher in all vignettes was White (National Center for Education Statistics, 2023).

Teachers were (truthfully) told that the main purpose of the study was to pilot the vignettes as an instrument to capture teachers' responses to classroom situations in math (see Hill et al., in press). We embedded a manipulation check at the end of the survey—a question asking teachers to report the race-ethnicity of the students featured in the vignettes—to test if race-ethnicity was adequately salient and operationalized. Overall, 87% of teachers correctly noted the race of the

student speaking—86%, 82%, and 92% in the classroom condition featuring a majority of Black, Latinx/e, and White students, respectively. We found emergent evidence to suggest teachers in the condition featuring mostly White students may have been more likely to correctly notice and report the race of the students in the vignettes ($\chi^2=4.65, p=0.10$).

We acknowledge that student identities are complex and intersectional, and that race and ethnicity cannot be fully disentangled. While the categories we use are certainly not necessarily mutually exclusive in the real world, e.g. in the case of Afro-Latinx students, our purpose was to attempt to make the student's racial-ethnic identity salient in order to measure teacher biases. We accept this trade-off between the study's simplification of student race-ethnicity and the benefit of an experimental design.

Procedure

Vignettes were presented to teachers via the survey software Qualtrics. We fixed the first vignette to randomize teachers into their condition, and the subsequent five were presented in random order to avoid potential order effects. After respondents viewed the frames depicting the sequence of instruction in each vignette, the final frame demonstrated the target student response(s) and directed participants to “Speak what you would say next to the class” as if they were the hypothetical teacher (Question 1, Q1). Next, they were asked a second question: “After saying this to the class, what would you do next?” (Question 2, Q2). We captured teacher responses using an embedded voice transcription software, *Phonic*.

After recording their responses to the open-ended Q1 and Q2 questions, teachers responded to closed-ended Likert-type items pertaining to each vignette using a 1 (low) - 5 (high) scale. The survey concluded with questions limited to teacher background. In addition to attention checks, we

included manipulation checks throughout the survey to determine whether respondents noticed the race/ethnicity of the students in the vignettes.

Measures

See Table 1.3 for a summary of the prompts and nature of data for all pre-registered measures. We include, for completeness, our measure of eliciting student thinking, which we were unable to analyze because of poor fit between the items and construct.

Table 1.3. Constructs Measured and Prompts in the Survey

<i>Construct</i>	<i>Survey Prompts</i>
Teacher Judgments About the Quality of Students' Mathematical Thinking (Likert, 1-5 scale)	How mathematically sophisticated is [student]'s response? How likely is it that [student] has a solid conceptual understanding of [problem type]? [Student]'s response indicates they are on track toward mastering [problem type]. [Student]'s thinking is mathematically sound. The teacher should encourage [student] to solve similar problems in the same way. The teacher should encourage the class to solve similar problems in the same way.
Task Difficulty (1-5 scale)	How difficult would this problem be for the typical 4th grader in the U.S.?
Eliciting Student Thinking (EST)	Q1: Speak what you would say next to the class.
Using Student Thinking (UST)	Q1: Speak what you would say next to the class.
Teacher Affirmation (TA)	Q1: Speak what you would say next to the class.
Supporting Student Sense-Making (SSS)	Q2: After saying this to the class, what would you do next?

Teacher Expectations and Judgments About Students' Math Abilities

Teacher Judgments about the Quality of Student Mathematical Thinking. We measured teacher judgments about the quality of the student's mathematical thinking via six Likert-type items (1-5 scale) per vignette where the teacher rated the student response's sophistication, the student's likely

level of understanding, their content mastery, the soundness of their thinking, and the extent to which the teacher would encourage the student and class to use the approach described by the target student (an endorsement of the student’s strategy). The items were developed based on prior scholarship asking teachers to evaluate students’ written mathematical work along dimensions of accuracy/correctness, mathematical clarity, sophistication of solution, mathematical language, and written mechanics and which found the items to load onto a single latent factor (Pollard, 2022). This measure also aligns with prior research on teachers’ evaluations of students’ mathematical thinking—which captured teachers’ ratings for the quality of student explanations, representations, understanding, sophistication, and mathematical competence—and composed them into a single measure and construct (Battey et al., 2021). Our measure had acceptable internal reliability ($\alpha=0.90$) across the 36 items. We constructed and used scores output by a nested measurement model within vignette due to the hierarchical data structure, with the six ratings under each vignette, and a confirmatory factor analysis (CFA) revealed adequate fit ($\chi^2(580) = 1311.44, p < 0.00$ CFI = 0.925 TLI = 0.919 RMSEA = 0.06 SRMR = 0.07; Hu & Bentler, 1999). This model specification proved to be the best fit to the data when comparing the model fit of alternative specifications (e.g. nesting within item type rather than vignette). See Figure B1.1 for a conceptual diagram and CFA for this measure.

Teacher Expectations of Task Difficulty for the Student. We measured teacher expectations for task difficulty with one item per vignette intended to capture how difficult they expected the math problem to be for “the typical fourth grader.” Teachers responded on a 5-point Likert scale, with 1 meaning not difficult at all and 5 being very difficult. Teachers’ reported ratings of difficulty were cohesive across the six vignettes ($\alpha = 0.79$), and a confirmatory factor analysis revealed adequate fit ($\chi^2(9) = 17.09, p = 0.05$, CFI=0.995, TLI=0.991, RMSEA = 0.05, SRMR = 0.04; Hu & Bentler, 1999). See Figure B1.2 for a conceptual diagram and CFA for this measure.

Teacher Responsiveness to Student Thinking

Before analysis, all open-ended responses were transcribed and reviewed against their original audio by human raters; the ASR transcriptions were 95% accurate at the word level, and the meaning of most inaccurate words could be inferred. Members of the project team developed an initial set of low-inference codes for the teacher responsiveness measures and completed preliminary coding on pilot data to achieve 80% inter-rater agreement or better, refining these codes to improve clarity. In rare cases, we coded responses as N/A when there was no teacher utterance or it was not responsive to the prompt. However, we only retained teachers in our sample who had codable responses on at least four out of the six vignettes. Once codes were finalized, the lead team recruited and trained raters using scores from the pilot data. Raters specialized in a particular code and double-scored responses within each vignette until an average 80% inter-rater agreement rate or better was achieved. The first author held weekly norming sessions for the rating team to discuss and resolve discrepancies. Raters went on to single score responses, and a member of the project team double-scored 20% of those scores to assess rater drift and accuracy. The project team resolved any disagreements that occurred during this process. Table A1.1 in Appendix A1 provides codes for all teacher responsiveness measures.

Eliciting Student Thinking (EST). We coded teacher responses to Q1 for the extent to which teachers elicited target student ideas in each vignette. We assigned a score of zero when there was no elicitation of any additional information from the student (e.g. “Great work!”), a one for minimal or generic elicitation that did not draw on anything specific the student had shared (e.g. “Can you explain to everyone what you mean?”), and a two for strong elicitation when the teacher posed a question to clarify or probe a specific idea that the student had shared (e.g., “When does it make the most sense to use your taking-away-one-adding-one strategy?”). The inter-rater agreement rate was

0.90 for EST; however, while a factor analysis indicated a distinct factor for these items, the six vignettes together exhibited poor score reliability, with $\alpha=0.34$ over the six items. We thus excluded this construct from further analysis.

Using Student Thinking (UST). We coded teacher responses to Q1 for the extent to which teachers used student thinking in their subsequent utterance. We assigned a code of zero when the teacher responded in a pro forma way without using student thinking (e.g. “Who else would like to share?”); a one when they used some aspect of the student’s thinking, such as brief restatement; and a two when there was strong use of student thinking, such as the teacher weaving the student idea into the development of the mathematics. The interrater agreement for UST was 0.78, and we estimated $\alpha=0.72$ over the six items, demonstrating adequate internal consistency.

Supporting Student Sensemaking (SSS). We coded teacher responses to Q2 for supporting student sense-making, assigning a code of zero when there was no support for student sensemaking in the next planned instructional step (e.g. “Let’s move on to a new problem”), and a one when the teacher centered student sense-making (e.g. “Who agrees with the answer? Explain why or why not”). The interrater agreement rate for SSS was 0.84, and we estimated $\alpha=0.72$ over the six items, demonstrating adequate internal consistency.

After scoring for the three constructs above, we conducted exploratory factor analysis that permitted these scores to load onto all three hypothesized latent factors in an unrestricted measurement model. This analysis revealed the strongest standardized factor loadings generally corresponded with their respective hypothesized factors, suggesting a distinct factor solution for each respective construct (Hill et al., in press).

Teacher Affirmation (TA). We coded Q1 responses for teacher affirmation, assigning a code of zero when there was no affirmation (e.g. “Who agrees?”), a one for a general affirmation (e.g.

“Excellent job on that problem”), and a two for specific affirmation (e.g. “Great work explaining there are ten tens hidden in each hundred”). Rater agreement was 0.95 for teacher affirmation (TA). We conducted a separate factor analysis for this affective construct. For the six vignette scores for teacher affirmation in Q1 (responding to the target student directly), we fit an adequate measurement model ($\chi^2(9) = 16.00, p = 0.07, RMSEA = 0.05, CFI = 0.99, TLI = 0.98, SRMR = 0.04, WRMR = 0.59$). Teachers’ coded responses for TA were also cohesive ($\alpha = 0.77$). These findings support a distinct factor solution for teacher affirmation.

For all coded measures of teacher responsiveness that were adequately internally consistent—UST, SSS, and TA—we estimated scores for each teacher using 2-parameter logistic (2PL) models with Item Response Theory’s (IRT) Graded Response Model (GRM) for polytomous data (Samejima, 1969).

Other Measures

Teacher background. We use an ordinal variable to measure teaching experience: less than 1 year (0), 1-2 years (1), 3-5 years (2), 6-10 years (3), 11-15 years (4), 16-20 years (5), and >20 years (6). For grade level, we also use an ordinal variable: Pre-K to grade 2 (0), mid-elementary or grades 3-5 (1), middle school or grades 6-8 (2), and high school or grades 9-12 (3). For respondents who indicated multiple grade levels, we used their lowest level taught to obtain a conservative proxy for level of teacher mathematical curricular knowledge. Teacher gender was measured as woman, man, or non-binary. Finally, we used dummy variables to represent the following self-reported teacher identities: Black, Latinx/e, White, Asian American or Pacific Islander, Native American, or multiracial. Due to small sample sizes for some of these identities, we could not conduct interaction analyses between condition and teacher race. These small sample sizes and a lack of theory also prevented us from interpreting the coefficients for teacher race-ethnicity in our results below.

Teaching context covariates. Some have argued that racial biases in teacher beliefs and behaviors may simply reflect their experiences with racial disparities in student achievement (Jussim & Harber, 2005). Teacher expectations are informed by prior student achievement, socioeconomic status, gender, student diagnostic labeling, and student race-ethnicity (Peterson et al., 2016). As such, we asked teachers for their school zip code and cross-referenced that with the Stanford Education Data Archive (SEDA) to obtain the following covariates: school urbanicity as measured by the proportion of students in city or urban locale schools, the percentage of White students in the locale, a composite score describing the metro locale's socioeconomic status, and finally a cohort scale mean of the math achievement in the metro area (Fahle et al., 2021).

Analysis

Due to the experimental nature of the study, we preregistered on the Open Science Framework (OSF) for scholarly transparency. We investigated the potential for bias in teachers' scores on the above dimensions by fitting a series of multiple regression models with student race-ethnicity entered as a fixed effect, with the reference group being the White classroom condition. We used structural equation modeling (SEM) with maximum likelihood and diagonally weighted least squares estimation, respectively, to model the impact of condition on teacher judgments about student mathematical thinking and expectation of math task difficulty, as this technique allowed us to account for measurement error in the construction of our latent variables, allowing us to be more precise compared to a simple regression (Kline, 2016). See Figures B1.3 and B1.4 for the analytic diagrams of these two SEM models. We modeled the impact of condition on our responsiveness measures by running a series of individual multiple regressions with teachers' estimated theta scores as the outcome. To improve the precision of our estimates, we include the demographic and covariates variables described earlier in both the SEM and multiple regression models (Equation 1).

$$\textbf{Equation 1} \quad Y_j = B_{0j} + B_1\textit{Black}_j + B_2\textit{Latinx}_j + B_3X_j + \textit{error}_j$$

Where j =participant, X represents the set of covariates described above, and error describes a random error variance term. B_1 and B_2 are the coefficients measuring the main effects of bias. A retrospective power analysis revealed we were between a power estimate of 0.17-0.49 to find the effects we found, which we take as substantively meaningful differences in teacher responses to students. Descriptive analyses and regressions were completed using Stata version 17.0, and factor analysis was completed using RStudio version 2023.12.1+402 with the lavaan package for robust model fit indicators to accommodate polytomous response data.

Results

Teacher Expectations and Judgements About Student Math Abilities

We did not find statistically significant differences by condition in teacher judgments about the quality of students' mathematical thinking (see Table 1.4). For teacher expectations, our findings show respondents expected greater task difficulty for the typical U.S. student on average when assigned to the Black ($\beta = 0.11, p=.046, \beta_{\text{std}} = 0.19$) and Latinx/e ($\beta = 0.14, p=.017, \beta_{\text{std}} = 0.23$) conditions, $\chi^2(79)= 74.70, p=.616, \text{CFI}=1.00, \text{TLI}=1.00, \text{RMSEA}=0.00, \text{SRMR} = 0.04$. Our findings also showed that those teaching more advanced grade levels on average expected greater task difficulty for students ($\beta = 0.14, p<.001, \beta_{\text{std}} = 0.24$). We also found that those with more teaching experience ($\beta = -0.05, p=.001, \beta_{\text{std}} = -0.08$) and those teaching in contexts with higher aggregate math achievement ($\beta = -0.38, p=.006, \beta_{\text{std}} = -0.62$) expected lower task difficulty for students, on average. Finally, we find trend-level, though not statistically significant, evidence to suggest teachers who are women may expect a lower level of difficulty (i.e. higher expectations) for students overall ($\beta = -0.12, p=.094, \beta_{\text{std}} = -0.20$) compared to teachers who are men.

Teacher Responsiveness to Student Mathematical Explanations

We did not observe statistically significant differences by condition in teacher use of student thinking. Those with more years of teaching experience were more likely to use student thinking ($\beta = 0.07, p=.027$). We reiterate that our sample was insufficiently powered for interpretation of the coefficients for teacher racial-ethnic subgroups. When articulating their next instructional steps, we found no statistically significant differences by condition, though we saw a trend suggesting teachers may have been more likely to support student sensemaking for Black students relative to White students ($\beta = 0.20, p=.095$). Women were also more likely to support student sensemaking compared to men ($\beta = 0.31, p=.030$). Finally, we did not find statistically significant differences in teacher affirmations across conditions, though we observed a trend suggesting teachers may have provided a greater degree of substantive affirmations to Black students relative to White students for the same mathematical solutions ($\beta = 0.21, p=.101$).

Table 1.4. Results from SEM and Multiple Regression Models

	Judgments About the Quality of Student Thinking		Expectation of Task Difficulty		Use of Student Thinking (UST)		Supporting Student Sensemaking (SSS)		Affirmation (TA)	
Black Classroom	0.07 (0.18)	0.09 (0.19)	0.16** (0.06)	0.11* (0.06)	0.17 (0.12)	0.13 (0.12)	0.24* (0.11)	0.20† (0.12)	0.26* (0.12)	0.21† (0.13)
Latinx/e Classroom	0.07 (0.18)	0.07 (0.19)	0.13* (0.06)	0.14* (0.06)	0.12 (0.12)	0.12 (0.12)	0.23† (0.12)	0.17 (0.12)	0.01 (0.12)	-0.03 (0.13)
Woman		0.18 (0.23)		-0.12† (0.07)		-0.04 (0.15)		0.31* (0.14)		0.02 (0.15)
Years Teaching		0.00 (0.05)		-0.05** (0.01)		0.07* (0.03)		0.05 (0.03)		-0.02 (0.03)
Grade level		-0.04 (0.09)		0.14*** (0.03)		-0.04 (0.06)		0.01 (0.06)		0.05 (0.06)

Table 1.4. Continued

	Judgments About the Quality of Student Thinking	Expectation of Task Difficulty	Use of Student Thinking (UST)	Supporting Student Sensemaking (SSS)	Affirmation (TA)			
Black Teacher	0.20 (0.31)	0.14 (0.09)	-0.36† (0.20)	-0.35† (0.19)	-0.30 (0.21)			
Latinx/e Teacher	-0.01 (0.35)	0.00 (0.11)	0.26 (0.23)	0.38† (0.22)	0.11 (0.24)			
AAPI Teacher	-0.30 (0.47)	-0.20 (0.13)	0.15 (0.27)	0.51† (0.27)	0.00 (0.29)			
Native Am. Teacher	0.57 (0.77)	-0.12 (0.30)	-1.37** (0.50)	-0.58 (0.48)	-0.24 (0.53)			
Multiracial Teacher	-0.07 (0.31)	-0.07 (0.10)	-0.04 (0.20)	-0.22 (0.20)	-0.21 (0.21)			
Urbanicity	0.49 (0.45)	-0.16 (0.14)	0.14 (0.29)	-0.10 (0.28)	0.09 (0.31)			
% White students	-0.29 (0.47)	-0.03 (0.14)	0.02 (0.30)	0.24 (0.29)	-0.15 (0.32)			
Metro Area SES	0.07 (0.18)	0.04 (0.05)	0.03 (0.11)	0.10 (0.11)	0.02 (0.12)			
Math Achievement	0.20 (0.44)	-0.38** (0.14)	-0.11 (0.29)	-0.30 (0.28)	-0.08 (0.30)			
			R ² = 0.01	R ² = 0.07	R ² = 0.02	R ² = 0.09	R ² = 0.02	R ² = 0.04

*** $p < .001$, ** $p < 0.01$, * $p < 0.05$, † ≤ 0.10

Note: Coefficients are shown with standard errors in parentheses. $N=312$ for all unconditional models and $N=307$ for all conditional models due to five observations missing covariate data. AAPI stands for Asian American and Pacific Islander. Metro area SES is a composite score describing the socioeconomic status of the metro locale in which the teacher works. For our two scale measures (judgments and expectations), we interpret coefficients on a 1-5 scale. For our coded responses (UST, SSS, and TA), we interpret coefficients as standardized differences in the latent construct, θ .

Discussion

In this study, we replicate and extend the literature about biases in teacher expectations to show that teachers' expectations of math task difficulty for the typical student, but not their judgments about the quality of students' mathematical thinking, may be biased against minoritized students of color. We further contribute to the evidence base on biases in education to include verbal teacher responsiveness to correct and non-standard student math solutions, finding no statistically significant differences by condition in dimensions of teacher responsiveness. However, we find a trend suggesting study participants may be more likely to affirm and support the sense-making of target students in the condition featuring a majority of Black students. We situate each of our findings in the existing literature below, then consider implications for research and practice.

Prior experimental work in mathematics showing incorrect and partially correct student work has demonstrated racial-ethnic biases against racially minoritized students in teacher expectations and judgments about student abilities, yet no biases in their evaluation of the student work itself (Pollard, 2022; Copur-Gencturk et al., 2019). We extend these experimental findings to student work free of errors showing non-standard and innovative problem-solving approaches. Taken together, these findings suggest that when evaluating and judging the quality of the mathematical work itself—whether the work shows errors or correct or even innovative solutions—teachers tend to not exhibit racial ethnic biases. Given that our findings contrast with findings from Battey et al., (2021) which identified biases against Black students in teacher judgements of student thinking in the context of mathematically correct student work presented via video, our findings suggest that race-ethnicity alone may not necessarily trigger teacher biases, but its interactions with students' sociocultural ways of speaking about math might instead be at play (e.g. Nasir et al., 2008; Shah, 2021).

When the focus shifts to evaluations and judgments about the typical *student*, we see a different picture. Copur-Gencturk and colleagues (2019) found teachers' math ability ratings for Black and Latinx/e girls to be the lowest compared to boys and White students after viewing their partially correct math solutions. We extend these findings to teachers' expectations of difficulty after viewing correct and non-standard student solutions. Our findings show teachers considered math tasks more difficult for the typical student when presented with 'classrooms' primarily composed of Black and Latinx/e students as opposed to White students, even when the student's mathematical work was correct and even innovative in its approach. An alternative explanation for this finding could be teachers accurately assessing problem difficulty based on pre-existing racialized achievement disparities, (e.g. Jussim & Harber, 2005). Still, because prior scholarship has identified biases above and beyond student achievement (Peterson et al., 2016) and the content of our vignettes was identical across conditions save for the race-ethnicity of the students, we argue stereotyped associations shaped biases in teacher perceptions of difficulty rather than social reality. Because teacher expectations relate to curricular choices (Aydin & Ok, 2022), this finding may be consequential for the curriculum students receive. For instance, expectations about task difficulty may underlie the biases against Black students in teachers' curricular selection in mathematics found in Pollard (2022). Such differences in expectations may contribute to noted differences in the intended versus enacted curriculum students experience in mathematics, for example, by task routinization, remediation or simplification (Stein et al., 2007; TNTP, 2018). As our findings showed teachers expect greater difficulty for Black and Latinx/e learners based on the same and correct mathematical content, our study raises questions about how teachers attribute those difficulties and what kinds of instructional experiences follow, with important implications for student opportunity to learn.

This study also extends prior findings on biases in instruction to the domain of responsive teaching in math. We did not find statistically significant differences by condition in teachers' use of student thinking, support of student sensemaking, and affirmations, though our findings showed trends indicating teachers may be more likely to support student sensemaking in their next instructional steps and may provide more positive and substantive affirmations for Black students relative to White students for the same mathematical content. While we interpret these findings with caution, they echo prior research by Harber et al. (2012) and Kleen and Glock (2018), which found teachers demonstrate positive feedback bias toward Black, Latinx/e, and ethnic minority students. The shifting standards theory may help to make sense of these findings which may run counter to hypotheses about biases, which states that stereotypes about social groups inform standards for evaluating members of that group, and those standards shift and differ across groups (Biernat et al., 1991; Biernat, 1995). Thus, our emergent findings may result from teachers showing a positive bias for minoritized students who excel academically. In similar work, Kleen and Glock (2018) found teachers judged high-performing ethnic minority students more favorably than high-performing ethnic majority students.

Fiske & Neuberg's (1990) dual-process theory for arriving at social judgments may help to interpret our seemingly contradictory findings. As our intervention presented teachers in all conditions with a student's proficient mathematical thinking and primed them to respond, teachers may have been confronted with information that contradicted activated stereotypes about learners. Our results showed teachers expected a greater degree of difficulty for students of color, yet they evaluated the quality of and used students' mathematical thinking comparably across conditions. Here, teacher perceptions about student proficiency may have contradicted any stereotypes, mitigating the potential for bias. When presented with information about a person that contradicts a stereotype, an information-integrating strategy is employed whereby individual attributes are

processed in an effortful, conscious manner to arrive at a judgment (Fiske & Neuburg, 1990). Such effortful and controlled processing may have worked to counteract perceived stereotypes (Pendry & Macrae, 1994), resulting in relatively equitable evaluations and use of student thinking across groups.

To summarize, our findings reproduce prior experimental findings regarding teacher expectations and judgments about student ability and extend them to contexts featuring correct, non-standard and even innovative student math solutions. We find that when evaluating and judging the quality of students' mathematical work itself, it is unlikely that teachers will exhibit biases, but when the focus shifts to evaluations and judgments about the student, teachers expect Black and Latinx/e students to experience greater difficulties. Second, we extend prior work about biases in instruction to the domain of responsive math teaching, finding suggestive rather than conclusive evidence in need of further investigation in the field. Finally, we extend previous work on racial-ethnic biases in the domain of mathematics to include Latinx/e students in addition to Black students.

Limitations

We reflect on several limitations to the present study. To start, we may have inadvertently intervened on what we intended to measure: it could be the case that we did not observe biases in some domains of responsive math instruction because we prompted all teachers in our study to respond to the student's mathematical thinking. An understanding of teachers' responses in the classroom also includes *not* responding (Kang, 2022) or *not* noticing students' mathematical contributions, particularly those from racially minoritized learners (Louie, 2018). It remains to be observed whether educators would have chosen to respond to the student's mathematical ideas had they not been universally prompted to do so and what the consequences would be. Prompting

teachers to take up student thinking could have played a leveling role in student opportunities to learn.

Another limitation concerns the demographic make-up of our sample. We did not have enough power to enable an examination of interaction effects between teacher and student race-ethnicity. Our sample consisted primarily of White educators, and White teachers may demonstrate social desirability bias in efforts to maintain a positive self-image and not appear biased or racist (Crosby & Monin, 2007; Plant & Devine, 1998). Moreover, even though they only composed 7% of our sample, Black teachers were somewhat overrepresented in the condition featuring White students, and White teachers may have been slightly overrepresented in the condition featuring Black students. There is ample literature on the positive effects of racial-ethnic congruence or “match” between teachers and students, for example when Black students are taught by Black teachers (Gershenson et al., 2016; Redding, 2019) and the nature of relational interactions for racially (mis)matched teachers and students in middle school mathematics (Battey et al., 2018). Had more Black teachers and fewer White teachers been in the condition featuring Black students, we may have found higher expectations and less or no bias overall. Similarly, our sample leaned towards experienced educators, which have been shown to be less biased than novice teachers (Vinopal & Holt, 2019); had our sample consisted of more early-career teachers, we may have found more pronounced biases against Black and Latinx/e learners. Future work investigating racial biases in responsive teaching should include a stratified, racially diverse sample of teachers with more varied levels of experience and attend to the phenomenon of racial-ethnic match between teacher and student.

A final limitation stems from the study’s design, which precluded us from three important lines of inquiry. First, due to sample size and power limitations, we could not take an intersectional

approach that examines interactions between student gender, race-ethnicity, and socioeconomic status. Biases are complex and are likely influenced by the intersections between various student and teacher identity factors (e.g. Battey et al., 2018; Riegle-Crumb & Humphries, 2012). Next, the relatively low R^2 values in our models indicate that there are other variables beyond those we investigated that can better account for the variation in teacher responsiveness to students. For example, future research examining biases in teacher responsiveness might account for teachers' mathematical content and pedagogical knowledge, which we would expect to predict one's propensity to teach responsively (e.g. Mathematical Knowledge for Teaching, or MKT as investigated in relation to biases by Battey et al., 2021 and Hill et al., 2008).

A final limitation from the study design concerns the format and content of the vignettes. In terms of format, the vignettes depicted a decontextualized, hypothetical instructional scenario which may not capture teachers' responsiveness in practice with real students with whom they have existing relationships. Given that improved empathy for and relationships with students can reduce racial biases in teacher behavior (e.g. Okonafua et al., 2016), future work must examine the extent to which any biases in responsive teaching may be mediated by relationships. Additionally, while prior research has linked teacher responses to vignettes to their classroom practices (e.g. Stecher et al., 2006 ; Stein et al., 2017), more external validity work is needed to connect teacher responses to instructional vignettes to their classroom practice, as teacher responsiveness may differ across both settings. In terms of vignette content, the math problems were at the fourth-grade level; and while we controlled for teachers' grade level in our model, those working in younger or more advanced grades may have differing expectations of difficulty in terms of problem content alone. Additionally, all student answers and solutions featured in the vignettes presented were mathematically correct, with some even demonstrating student reasoning through non-standard problem-solving strategies. This context may have made it clear to teachers that students across groups were equally proficient

in the particular mathematics topics represented in the vignettes. This is further supported by the fact that we did not detect any differences in teachers' use and judgments about the quality of student thinking. Such correct solutions may have provided stereotype-disconfirming evidence to teachers, reducing the potential influence of racial biases. Student demonstration of mathematical proficiency is consequential for the instructional character of a classroom. Indeed, prior research has shown teachers to employ ambitious teaching practices (Lampert et al., 2013)—of which responsiveness from the teacher is a large part—in math less frequently the more students they had performing below grade level (Gottfried et al., 2023). Future work in this area should examine the role of student math proficiency in teachers' racial-ethnic biases in this domain.

Conclusions

The mixed findings from this study have implications for pre- and in-service teacher education as well as for further research on racial-ethnic biases in responsive math teaching. We discuss several of those implications. First, our findings suggest surfacing students' non-standard and sophisticated mathematical strategies may serve as stereotype disconfirming information. Providing opportunities for students to share their mathematical reasoning and problem-solving processes may serve as an equity practice. Second, our findings indicate a need to address biased estimations of difficulty against Black and Latinx/e students in mathematics. We hope our findings encourage trainings that focus on exposing in-service and pre-service teachers to counterstereotypic exemplars to reduce implicit racial biases (Lai et al., 2014). Further, our inconclusive findings on positive feedback bias in responsive teaching practices invite further research on the extent and nature of any biases and whether they may be shaped by or interact with other student identity factors, like language and culture (e.g. Shah, 2021). Knowledge about the specific instructional sub-domains and circumstances subject to biases can better equip teacher educators and practitioners to

intervene on them, for example, through practice-in-context (Mancenido et al., 2025) and coaching (Cohen et al., 2024).

Youth in U.S. schools are more racially and ethnically diverse than they have ever been, meanwhile the teaching profession remains predominantly White (National Center for Education Statistics, 2023). Such demographic change is not inconsequential, as White educators in the U.S. have been found to carry negative stereotypes and social distance from minoritized groups (Quinn & Stewart, 2019), and U.S. counties with higher aggregate teacher racial bias tend to show greater achievement disparities between Black and White students (Chin et al., 2020). Responsiveness to student thinking is characteristic of effective learning environments for children, so much so that there have been concerted efforts to rehaul mathematics instruction in this way, particularly in urban schools serving Black and Latinx/e communities (Anderson, 2014). Our findings indicate the wide scale implementation of math instructional reforms absent a consideration of race could undermine goals for equitable and just math teaching.

References

- Alexander, R. J. (2008). *Towards dialogic teaching: Rethinking classroom talk* (4th ed.). Dialogos.
- Anderson, C. R. (2014). Place matters: Mathematics education reform in urban schools. *Journal of Urban Mathematics Education*, 7(1), 7. <https://doi.org/10.21423/jume-v7i1a231>
- Anderson-Clark, T. N., Green, R. J., & Henley, T. B. (2008). The relationship between first names and teacher expectations for achievement motivation. *Journal of Language and Social Psychology*, 27(1), 94–99.
- Aydin, Ö., & Ok, A. (2022). A systematic review on teacher’s expectations and classroom behaviors. *International Journal of Curriculum and Instructional Studies*, 12(1), 247-274.
- Ball, D. (2018). “Presidential Address,” American Educational Research Association Annual Conference, New York.
- Battey, D., Bartell, T., Webel, C., & Lowry, A. (2021). Understanding the impact of racial attitudes on preservice teachers’ perceptions of children’s mathematical thinking. *Journal for Research in Mathematics Education*, 52(1), 62-93.
- Battey, D., Leyva, L. A., Williams, I., Belizario, V. A., Greco, R., & Shah, R. (2018). Racial (mis) match in middle school mathematics classrooms: Relational interactions as a racialized mechanism. *Harvard Educational Review*, 88(4), 455-482.
- Benson, T. A. & Fiarman, S. E. (2020). *Unconscious Bias in Schools : A Developmental Approach to Exploring Race and Racism, Revised Edition: Vol. Revised edition*. Harvard Education Press.
- Bertrand, M., & Duflo, E. (2016). Field experiments on discrimination. In A. V. Banerjee & E. Duflo (Eds.), *Handbook of economic field experiments* (Vol. 1, pp. 309–393). Amsterdam: North Holland Publishing.
- Biernat, M. (1995). The shifting standards model: Implications of stereotype accuracy for social judgment. In Y. T. Lee, L. J. Jussim, & C. R. McCauley (Eds.), *Stereotype accuracy: Toward appreciating group differences* (pp. 87–114). American Psychological Association.
- Biernat, M., Manis, M., & Nelson, T. E. (1991). Stereotypes and standards of judgment. *Journal of Personality and Social Psychology*, 60, 485–499. <https://doi.org/10.1037/0022-3514.60.4.485>
- Bonefeld, M., Dickhäuser, O. & Karst, K. (2020). Do preservice teachers’ judgments and judgment accuracy depend on students’ characteristics? The effect of gender and immigration background. *Social Psychology of Education*, 23(2), 189–216. <https://doi-org.ezp-prod1.hul.harvard.edu/10.1007/s11218-019-09533-2>
- Bondurant, L., & Reinholz, D. (2023). “Rahul is a math nerd” and “Mia can be a drama queen”: How mixed-reality simulations can perpetuate racist and sexist stereotypes. *Mathematics Teacher Educator*, 11(3), 189-209.
- Boston, M. (2012). Assessing instructional quality in mathematics. *The Elementary School Journal*, 113(1), 76–104. <https://doi.org/10.1086/666387>
- Brophy, J. E. (1981). Teacher praise: A functional analysis. *Review of educational research*, 51(1), 5-32.

- Brophy, J. E., & Good, T. L. (1974). *Teacher-student relationships: Causes and consequences*. Holt, Rinehart and Winston.
- Carpenter, T. P., Fennema, E., Peterson, P. L., Chiang, C. P., & Loef, M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: An experimental study. *American Educational Research Journal*, 26(4), 499-531. <https://doi.org/10.3102/00028312026004499>
- Charalambous, C. Y. (2020). Reflecting on the troubling relationship between teacher knowledge and instructional quality and making a case for using an animated teaching simulation to disentangle this relationship. *ZDM Mathematics Education*, 52(2), 219-240. <http://dx.doi.org/10.1007/s11858-019-01089-x>
- Chin, M. J., Quinn, D. M., Dhaliwal, T. K., Lovison, V. S. (2020). Bias in the air: A nationwide exploration of teachers' implicit racial attitudes, aggregate bias, and student outcomes. *Educational Researcher*, 49(8):566-578. <https://doi.org/10.3102/0013189X20937240>
- Clotfelter, C., Ladd, H. F., Clifton, C. R. (2023). Racial Differences in Student Access to High-Quality Teachers. *Education Finance and Policy*, 18(4): 738–752. https://doi.org/10.1162/edfp_a_00402
- Cohen, D. K., & Hill, H. C. (2000). Instructional Policy and Classroom Performance: The Mathematics Reform in California. *Teachers College Record: The Voice of Scholarship in Education*, 102(2), 294-343. <https://doi.org/10.1111/0161-4681.00057>
- Cohen, J., Wong, V. C., Krishnamachari, A., & Erickson, S. (2024). Experimental evidence on the robustness of coaching supports in teacher education. *Educational Researcher*, 53(1), 19-35. <https://doi.org/10.3102/0013189X231198827>
- Copur-Gencturk, Y., Cimpian, J. R., Lubienski, S. T., & Thacker, I. (2019). Teachers' bias against the mathematical ability of female, Black, and Hispanic students. *Educational Researcher*, 49(1), 30-43. <https://doi.org/10.3102/0013189X19890577>
- Copur-Gencturk, Y., Thacker, I., & Cimpian, J. R. (2022). Teacher bias in the virtual classroom. *Computers & Education*, 191, 104627. <https://doi.org/10.1016/j.compedu.2022.104627>
- Copur-Gencturk, Y., Thacker, I., & Cimpian, J. R. (2023). Teachers' race and gender biases and the moderating effects of their beliefs and dispositions. *International Journal of STEM Education*, 10(1), 31. <https://doi.org/10.1186/s40594-023-00420-z>
- Crosby, J. R., & Monin, B. (2007). Failure to warn: How student race affects warnings of potential academic difficulty. *Journal of Experimental Social Psychology*, 43(4), 663–670. <https://doi.org/10.1016/j.jesp.2006.06.007>
- Cross, W. E. Jr., & Cross, T. B. (2008). Theory, research, and models. In Quintana, S. M. & McKown, C (Eds.), *Handbook of race, racism, and the developing child* (pp. 154–181). Wiley.
- Datnow, A., & Hubbard, L. (2015). Teachers' use of assessment data to inform instruction: Lessons from the past and prospects for the future. *Teachers College Record*, 117(4), 1-26.
- Davis, J. (2014). The Mathematical Experiences of Black Males in a Predominantly Black Urban Middle School and Community. *International Journal of Education in Mathematics, Science and Technology*, 2(3), 206–222. <https://doi.org/10.18404/ijemst.55341>

- Dovidio, J. F. (2001). On the Nature of Contemporary Prejudice: The Third Wave. *Journal of Social Issues*, 57(4), 829–849. <https://doi.org/10.1111/0022-4537.00244>
- Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Harcourt Brace Jovanovich College Publishers.
- Ewing, E. L. (2020, July 2). I'm a Black scholar who studies race. here's why I capitalize 'White.' ZORA. <https://zora.medium.com/im-a-black-scholar-who-studies-race-here-s-why-i-capitalize-white-f94883aa2dd3>
- Fahle, E. M., Chavez, B., Kalogrides, D., Shear, B. R., Reardon, S. F., & Ho, A. D. (2021). *Stanford Education Data Archive: Technical Documentation (Version 4.1)*. Retrieved from <http://purl.stanford.edu/db586ns4974>.
- Fazio, R. H. (1990). Multiple Processes by which Attitudes Guide Behavior: The Mode Model as an Integrative Framework. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 23, pp. 75–109). Academic Press. [https://doi.org/10.1016/S0065-2601\(08\)60318-4](https://doi.org/10.1016/S0065-2601(08)60318-4)
- Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, 11(2), 77-83. <https://doi.org/10.1016/j.tics.2006.11.005>
- Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 23, pp. 1–74). Academic Press. [https://doi.org/10.1016/S0065-2601\(08\)60317-2](https://doi.org/10.1016/S0065-2601(08)60317-2)
- Flores, A. (2007). Examining disparities in mathematics education: Achievement gap or opportunity gap? *The High School Journal*, 91(1), 29–42. <https://doi.org/10.1353/hsj.2007.0022>
- Franke, M. L., Webb, N. M., Chan, A. G., Ing, M., Freund, D., & Battey, D. (2009). Teacher Questioning to Elicit Students' Mathematical Thinking in Elementary School Classrooms. *Journal of Teacher Education*, 60(4), 380–392. <https://doi.org/10.1177/0022487109339906>
- Gaddis, S. M. (2017a). How Black Are Lakisha and Jamal? Racial Perceptions from Names Used in Correspondence Audit Studies. *Sociological Science*, 4(19), 469–489. <https://doi.org/10.15195/v4.a19>
- Gaddis, S. M. (2017b). Racial/Ethnic Perceptions from Hispanic Names: Selecting Names to Test for Discrimination. *Socius*, 3. <https://doi.org/10.1177/2378023117737193>
- Gershenson, S., Holt, S. B., & Papageorge, N. W. (2016). Who believes in me? The effect of student–teacher demographic match on teacher expectations. *Economics of Education Review*, 52, 209–224. <https://doi.org/10.1016/j.econedurev.2016.03.002>
- Gottfried, M., Fletcher, T., & Comstock, M. (2023). Teaching mathematics in kindergarten: How does instruction differ by classroom ability? *Teachers College Record*, 125(10), 102–130. <https://doi.org/10.1177/01614681231223343>
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, 102(1), 4–27. <https://doi.org/10.1037/0033-295X.102.1.4>

- Greenwald, A. G., Banaji, M. R., & Nosek, B. A. (2014). Supplemental Material for Predicting Ethnic and Racial Discrimination: A Meta-Analysis of IAT Criterion Studies. *Journal of Personality and Social Psychology*, supp. <https://doi.org/10.1037/a0032734.supp>
- Gross, S. (1993). Early Mathematics Performance and Achievement: Results of a Study Within a Large Suburban School System. *The Journal of Negro Education*, 62(3), 269–287. <https://doi.org/10.2307/2295465>
- Gutiérrez, R. (2013). The Sociopolitical Turn in Mathematics Education. *Journal for Research in Mathematics Education*, 44(1), 37–68. <https://doi.org/10.5951/jresmetheduc.44.1.0037>
- Gutiérrez, R. (2012). Context matters: How should we conceptualize equity in mathematics education? In B. Herbel-Eisenmann, J. Choppin, D. Wagner, & D. Pimm (Eds.), *Equity in discourse for mathematics education: Theories, practices, and policies* (pp. 17–33). Springer. https://doi.org/10.1007/978-94-007-2813-4_2
- Harber, K. D., Gorman, J. L., Gengaro, F. P., Butisingh, S., Tsang, W., & Ouellette, R. (2012). Students' race and teachers' social support affect the positive feedback bias in public schools. *Journal of Educational Psychology*, 104(4), 1149–1161. <https://doi.org/10.1037/a0028110>
- Herbel-Eisenmann, B. A., Steele, M. D., & Cirillo, M. (2013). (Developing) teacher discourse moves: A framework for professional development. *Mathematics Teacher Educator*, 1(2), 181–196.
- Herbst, P., Chazan, D., Chen, C. L., Chieu, V. M., & Weiss, M. (2011). Using comics-based representations of teaching, and technology, to bring practice to teacher education courses. *ZDM Mathematics Education*, 43, 91–103. <https://doi.org/10.1007/s11858-010-0290-5>
- Hill, H. C., Ball, D. L., & Schilling, S. G. (2008). Unpacking Pedagogical Content Knowledge: Conceptualizing and Measuring Teachers' Topic-Specific Knowledge of Students. *Journal for Research in Mathematics Education*, 39(4), 372–400. <http://www.jstor.org/stable/40539304>
- Hill, H. C. Garcia Coppersmith, J., Kleen, H., Pollard, C. (in press). Measuring Standards-Based Instructional Practices at Scale: A New Approach. *Educational Assessment*.
- Hilton, J. L., & Von Hippel, W. (1996). Stereotypes. *Annual Review of Psychology*, 47, 237–271. <https://doi.org/10.1146/annurev.psych.47.1.237>
- Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A Meta-Analysis on the Correlation Between the Implicit Association Test and Explicit Self-Report Measures. *Personality & Social Psychology Bulletin*, 31(10), 1369–1385. <https://doi.org/10.1177/0146167205275613>
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55.
- Irvine, J. J. (2010). Foreword. In Milner's H. R. (Ed.), *Culture, curriculum, and identity in education* (pp. xi–xvi). Palgrave Macmillan
- Jacobs, V. R., Lamb, L. L. C., & Philipp, R. A. (2010). Professional Noticing of Children's Mathematical Thinking. *Journal for Research in Mathematics Education*, 41(2), 169–202. <http://www.jstor.org/stable/20720130>

- Jacobs, V. R., & Empson, S. B. (2016). Responding to children's mathematical thinking in the moment: An emerging framework of teaching moves. *ZDM Mathematics Education*, 48, 185-197. <https://doi.org/10.1007/s11858-015-0717-0>
- Jencks, C. & Phillips, M. (Eds.). (2011). *The Black-White test score gap*. Brookings Institution Press,
- Jussim, L., & Harber, K. D. (2005). Teacher expectations and self-fulfilling prophecies: knowns and unknowns, resolved and unresolved controversies. *Personality and Social Psychology Review*, 9, 131-155. DOI:10.1207/s15327957pspr0902_3
- Kang, H. (2022). Teacher responsiveness that promotes equity in secondary science classrooms. *Cognition and Instruction*, 40(2), 206-232.
- Kleen, H. & Glock, S. (2018). A further look into ethnicity: The impact of stereotypical expectations on teachers' judgments of female ethnic minority students. *Social Psychology of Education*, 21(4), 759-773. <https://doi.org/10.1007/s11218-018-9451-0>
- Kline, R. B. (2016). *Principles and Practice of Structural Equation Modeling* (4th ed.). Guilford Press.
- Ladson-Billings, G. (2006). From the achievement gap to the education debt: Understanding achievement in U.S. schools. *Educational Researcher*, 35(7), 3-12. <https://doi.org/10.3102/0013189X035007003>
- Lai, C. K., Marini, M., Lehr, S. A., Cerruti, C., Shin, J. L., Joy-Gaba, J. A., Ho, A. K., Teachman, B. A., Wojcik, S. P., Koleva, S. P., Frazier, R. S., Heiphetz, L., Chen, E., Turner, R. N., Haidt, J., Kesebir, S., Hawkins, C. B., Schaefer, H. S., Rubichi, S., Sartori, G., Dial, C. M., Sriram, N., Banaji, M. R., & Nosek, B. A. (2014). Reducing Implicit Racial Preferences: I. A Comparative Investigation of 17 Interventions. *Journal of Experimental Psychology. General*, 143(4), 1765-1785. <https://doi.org/10.1037/a0036260>
- Lampert, M. (1990). Connecting inventions with conventions. In L. P. Steffe & T. Wood (Eds.), *Transforming children's mathematics education* (pp. 253-265). Lawrence Erlbaum Associates, Inc.
- Lampert, M. (2001). *Teaching problems and the problems of teaching*. Yale University Press.
- Lampert, M., Franke, M. L., Kazemi, E., Ghouseini, H., Turrou, A. C., Beasley, H., Cunard, A., & Crowe, K. (2013). Keeping it complex: Using rehearsals to support novice teacher learning of ambitious teaching. *Journal of Teacher Education*, 64(3), 226-243. <https://doi.org/10.1177/0022487112473837>
- Laursen, S. L., Hassi, M. L., Kogan, M., & Weston, T. J. (2014). Benefits for women and men of inquiry-based learning in college mathematics: A multi-institution study. *Journal for Research in Mathematics Education*, 45(4), 406-418.
- Laws, M. (2020). Why we capitalize 'Black' (and not 'white'). *Columbia Journalism Review*. Retrieved August 9, 2024, from <https://www.cjr.org/analysis/capital-b-black-styleguide.php>
- Leyva, L. A., McNeill, R. T., Marshall, B. L., & Guzmán, O. A. (2021). "It Seems like They Purposefully Try to Make as Many Kids Drop": An Analysis of Logics and Mechanisms of Racial-Gendered Inequality in Introductory Mathematics Instruction. *The Journal of Higher Education (Columbus)*, 92(5), 784-814. <https://doi.org/10.1080/00221546.2021.1879586>

- Lightfoot, S. L. (1978). *Worlds apart: Relationships between families and schools*. Basic Books
- Louie, N. L. (2018). Culture and ideology in mathematics teacher noticing. *Educational Studies in Mathematics*, 97(1), 55–69. <https://doi.org/10.1007/s10649-017-9775-2>
- Malouff, J. M., & Thorsteinsson, E. B. (2016). Bias in grading: A meta-analysis of experimental research findings. *Australian Journal of Education*, 60, 245–256. <https://doi.org/10.1177/0004944116664618>
- Mancenido, Z., Hill, H. C., Garcia Coppersmith, J., Carter, H., Pollard, C., & Monschauer, C. (2025). Practice-Based Teacher Education Pedagogies Improve Responsiveness: Evidence from a Lab Experiment. *Journal of Research on Educational Effectiveness*, 1–25. <https://doi.org/10.1080/19345747.2025.2456716>
- Marshall, Y. (2020) (In)articulate while Black—AAIHS. *Black Perspectives*. Retrieved August 11, 2024, from <https://www.aaihs.org/inarticulate-while-black/>
- Martin, D. B. (2009). Liberating the production of knowledge about African American children and mathematics. In D. B. Martin (Ed.), *Mathematics teaching, learning, and liberation in the lives of Black children* (pp. 3–36). Routledge/Taylor & Francis Group.
- Milner, H. R. (2010). *Start where you are, but don't stay there: understanding diversity, opportunity gaps, and teaching in today's classrooms*. Harvard Education Press.
- Milner, H. R. (2012). Beyond a test score: Explaining opportunity gaps in educational practice. *Journal of Black Studies*, 43(6), 693–718. <https://doi-org.ezp-prod1.hul.harvard.edu/10.1177/0021934712442539>
- Nasir, N. S., Hand, V., & Taylor, E. V. (2008). Culture and mathematics in school: Boundaries between “cultural” and “domain” knowledge in the mathematics classroom and beyond. *Review of Research in Education*, 32, 187–240.
- National Center for Education Statistics. (2024). Racial/Ethnic Enrollment in Public Schools. *Condition of Education*. U.S. Department of Education, Institute of Education Sciences. <https://nces.ed.gov/programs/coe/indicator/cge>.
- National Center for Education Statistics. (2023). *Characteristics of public school teachers*. Condition of education. U.S. Department of Education, Institute of Education Sciences. <https://nces.ed.gov/programs/coe/indicator/clr>
- National Center for Education Statistics. (2019). *Trends in International Mathematics and Science Study (TIMSS)*. National Center for Education Statistics.
- National Council of Teachers of Mathematics (NCTM). (2000). *Principles and standards for school mathematics*. NCTM.
- National Council of Teachers of Mathematics (NCTM). (2014). *Principles to actions: ensuring mathematical success for all*. NCTM.
- National Governors Association Center for Best Practices & Council of Chief State School Officers (2010). *Common Core State Standards for Mathematics*. Washington, DC

- O'Connor, M. C., & Michaels, S. (1996). Shifting participant frameworks: Orchestrating thinking practices in group discussion. In D. Hicks (Ed.), *Discourse, learning, and schooling* (pp. 63-103). Cambridge University Press.
- Okonofua, J. A., Paunesku, D., & Walton, G. M. (2016). Brief intervention to encourage empathic discipline cuts suspension rates in half among adolescents. *Proceedings of the National Academy of Sciences*, 113(19), 5221–5226. <https://doi.org/10.1073/pnas.1523698113>
- Pendry, L. F., & Macrae, C. N. (1994). Stereotypes and mental life: The case of the motivated but thwarted tactician. *Journal of Experimental Social Psychology*, 30, 303–325. <https://doi.org/10.1006/jesp.1994.1015>
- Peterson, E. R., Rubie-Davies, C., Osborne, D., & Sibley, C. (2016). Teachers' explicit expectations and implicit prejudiced attitudes to educational achievement: Relations with student achievement and the ethnic achievement gap. *Learning and Instruction*, 42, 123–140. <https://doi.org/10.1016/j.learninstruc.2016.01.010>
- Pit-ten Cate, I. M., & Glock, S. (2023). A systematic review on teachers' stereotypical beliefs and expectations: Effects of the intersectionality of students' gender and ethnicity. In P. W. St. J. Watson, C. M. Rubie-Davies, & B. Ertl (Eds.), *Routledge international handbook of gender beliefs, stereotype threat and teacher expectations* (1st ed., pp. 238–250). Routledge/Taylor & Francis Group. <https://doi.org/10.4324/9781003275763-24>
- Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*, 75(3), 811–832. <https://doi.org/10.1037/0022-3514.75.3.811>
- Pollard, C. (2022). Teaching, Learning, and Race: Toward the Identification of Mechanisms Underlying Own- and Other-Race Teacher Effects. Doctoral dissertation, Harvard University Graduate School of Arts and Sciences.
- Quinn, D. M. (2020a). Experimental effects of “achievement gap” news reporting on viewers' racial stereotypes, inequality explanations, and inequality prioritization. *Educational Researcher*, 49(7), 482–492. <https://doi.org/10.3102/0013189X20932469>
- Quinn, D. M. (2020b). Experimental evidence on teachers' racial bias in student evaluation: the role of grading scales. *Educational Evaluation and Policy Analysis*, 42(3), 375–392. <https://doi.org/10.3102/0162373720932188>
- Quinn, D. M., & Stewart, A. M. (2019). Examining the racial attitudes of white pre-K–12 educators. *The Elementary School Journal*, 120(2), 272–299. <https://doi.org/10.1086/705899>
- Reardon, S. F., Cimpian, J., & Weathers, E. S. (2015). Patterns and trends in racial/ethnic and socioeconomic academic achievement gaps. In H. F. Ladd, & M. E. Goertz (Eds.), *Handbook of Research in Education Finance and Policy* (2nd ed.). Routledge.
- Redding, C. (2019). A teacher like me: A review of the effect of student–teacher racial/ethnic matching on teacher perceptions of students and student academic and behavioral outcomes. *Review of Educational Research*, 89(4), 499–535. <https://doi.org/10.3102/0034654319853545>
- Reinholz, D., Johnson, E., Andrews-Larson, C., Stone-Johnstone, A., Smith, J., Mullins, B., ... & Shah, N. (2022). When active learning is inequitable: Women's participation predicts gender

- inequities in mathematical performance. *Journal for Research in Mathematics Education*, 53(3), 204-226.
- Reinholz, D. L., & Wilhelm, A. G. (2024). Race-gender D/discourses in mathematics education:(Re)-producing inequitable participation patterns across a diverse, instructionally-advanced urban district. *Urban Education*, 59(8), 2519-2549.
- Richards, J., & Robertson, A. D. (2015). A review of the research on responsive teaching in science and mathematics. In *Responsive teaching in science and mathematics* (1st ed., pp. 36–55). Routledge/Taylor & Francis Group. <https://doi.org/10.4324/9781315689302-2>
- Riegle-Crumb, C., & Humphries, M. (2012). Exploring bias in math teachers' perceptions of students' ability by gender and race/ethnicity. *Gender & Society*, 26(2), 290-322.
- Rist, R. (1970). Student social class and teacher expectations: The self-fulfilling prophecy in ghetto education. *Harvard Educational Review*, 40(3), 411-451.
- Roberts, S. O. (2022). *Dealing with Diversity in Psychology: Science and Ideology*. OSF. <https://doi.org/10.31234/osf.io/xk4yu>
- Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom; teacher expectation and pupils' intellectual development*. Holt, Rinehart and Winston.
- Rubie-Davies, C. M., Flint, A., & McDonald, L.G. (2012). Teacher beliefs, teacher characteristics and school contextual factors: What are the relationships? *British Journal of Educational Psychology*, 82(2), 270–288. <https://doi.org/10.1111/j.2044-8279.2011.02025.x>
- Rubie-Davies, C., Hattie, J., & Hamilton, R. (2006). Expecting the best for students: Teacher expectations and academic outcomes. *British Journal of Educational Psychology*, 76(3), 429–444. <https://doi.org/10.1348/000709905X53589>
- Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika*, 34, 1–97.
- Schneider, D. J. (2004). *The psychology of stereotyping*. Guilford Press.
- Schoenfeld, A., Fink, H., Zuñiga-Ruiz, S., Huang, S., Wei, X., & Chirinda, B. (2023). *Helping students become powerful mathematics thinkers: Case studies of teaching for robust understanding*. <https://doi.org/10.4324/9781003375197>
- Shah, N. (2021). Racial equity and justice in teaching and teacher education: Progress, tensions, and open questions. Spencer Foundation.
- Shaughnessy, M., & Boerst, T. A. (2018). Uncovering the skills that preservice teachers bring to teacher education: The practice of eliciting a student's thinking. *Journal of Teacher Education*, 69(1), 40-55. <https://doi.org/10.1177/00224871177025>
- Smylie, M., & Wenzel, S. (2006). *Promoting instructional improvement: A strategic human resource management perspective*. Consortium on Chicago School Research.

- Stecher, B., Le, V. N., Hamilton, L., Ryan, G., Robyn, A., & Lockwood, J. R. (2006). Using structured classroom vignettes to measure instructional practices in mathematics. *Educational Evaluation and Policy Analysis*, 28(2), 101-130.
- Stein, M. K., Correnti, R., Moore, D., Russell, J. L., & Kelly, K. (2017). Using theory and measurement to sharpen conceptualizations of mathematics teaching in the common core era. *AERA Open*, 3(1), 2332858416680566.
- Stein, M. K., Remillard, J., Smith, M. S. (2007). How curriculum influences student learning. In F. K. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 319-370). Charlotte, NC: Information Age Publishing.
- Stein, M. K., Engle, R. A., Smith, M. S., & Hughes, E. K. (2008). Orchestrating productive mathematical discussions: Five practices for helping teachers move beyond show and tell. *Mathematical Thinking and Learning*, 10(4), 313–340.
<https://doi.org/10.1080/10986060802229675>
- Sun, K. L., & Ruef, J. L. (2023). Examining and conceptualizing the relationship between teacher praise and the co-construction of mathematical competence in classrooms. *The Journal of Mathematical Behavior*, 71(3), 101065. <https://doi.org/10.1016/j.jmathb.2023.101065>
- Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American*, 223(5), 96–103.
- Tenenbaum, H. R., & Ruck, M. D. (2007). Are teachers' expectations different for racial minority than for European American students? A meta-analysis. *Journal of Educational Psychology*, 99(2), 253–273. <https://doi.org/10.1037/0022-0663.99.2.253>
- The New Teacher Project (TNTP) (2018). *The opportunity myth*. New York, NY. Retrieved September 20, 2018, from https://tntp.org/assets/documents/TNTP_The-Opportunity-Myth_Web.pdf
- Trejos, J. (2024). What's in a "Latinx?": Considerations when utilizing pan-ethnic identifiers of Latinidad. *Columbia Social Work Review*, 22(1), 28-48.
<https://doi.org/10.52214/cswr.v22i1.12235>
- van den Bergh, L., Denessen, E., Hornstra, L., Voeten, M., & Holland, R. W. (2010). The implicit prejudiced attitudes of teachers: Relations to teacher expectations and the ethnic achievement gap. *American Educational Research Journal*, 47(2), 497–527.
<https://doi.org/10.3102/0002831209353594>
- Vinopal, K., & Holt, S. B. (2019). Rookie Mistakes: The Interplay of Teacher Experience and Racial Representation. *Educational Researcher*, 48(7), 421–437.
<https://doi.org/10.3102/0013189X19867699>
- Warikoo, N., Sinclair, S., Fei, J., & Jacoby-Senghor, D. (2016). Examining racial bias in education: A new approach. *Educational Researcher*, 45(9), 508–514.
<https://doi.org/10.3102/0013189X16683408>
- Weinstein, R. S. (2002). *Reaching higher: The power of expectations in schooling*. Harvard University Press.
<https://doi.org/10.4159/9780674045040>
- Yogeeswaran, K., Devos, T., & Nash, K. (2018). Understanding the nature, measurement, and utility of implicit intergroup biases. In Sibley, C. G., & Barlow, F. K., *The Cambridge handbook of the psychology of prejudice* (Concise student edition) (pp. 69-99). Cambridge University Press.

Appendix A1

Table A1.1 Coding Scheme for Teacher Responsiveness Measures

Eliciting Student Thinking (EST)			
N/A	0 - No eliciting of student thinking.	1 - Some elicitation of student thinking	2 - Strong elicitation of student thinking
No teacher utterance or teacher utterance is not responsive to prompt.	Teacher makes a statement, whether approval or correction. No expectation of student response. Asks a question to the whole of the class / not the initial respondent (e.g. <i>"What do others think?"; "Did anyone else have a different way?"</i>)	Teacher poses a simple proforma clarifying or probing question that does not draw specifically on what the student has shared (e.g. <i>"How did you get that?"; "Can you explain to everyone what you mean?"</i>)	Teacher poses a question to clarify or probe a specific idea that the student has shared.
Using Student Thinking (UST)			
N/A	0 - No or minimal use of student thinking	1 - Some use of student thinking	2 - Strong use of student thinking
No teacher utterance or teacher utterance is not responsive to prompt.	Response is plausible in the given situation (e.g. <i>"Ok, nice!"</i>). Teacher responds in a pro forma way (e.g., acknowledges student response is correct or incorrect; thanks or praises the student but does not respond to their mathematical ideas; asks for another student answer to the problem; provides direct instruction about the problem to the class).	Teacher response goes beyond pro forma to feature some use of student ideas; focus stays at least for a moment on the student(s) but it does not rise to strong use (e.g., brief restatement of student method; brief restatement followed by teacher direct instruction; asking students to repeat their answers; asks class whether student is correct).	Teacher weaves student ideas into the development of the mathematics (e.g., compares student responses; adds mathematical emphasis to student method or solution; fills in mathematical details that were missing from student response; asks student "why" question; asks class "why" question based on a student method; asks <i>"How do you know?"</i>)

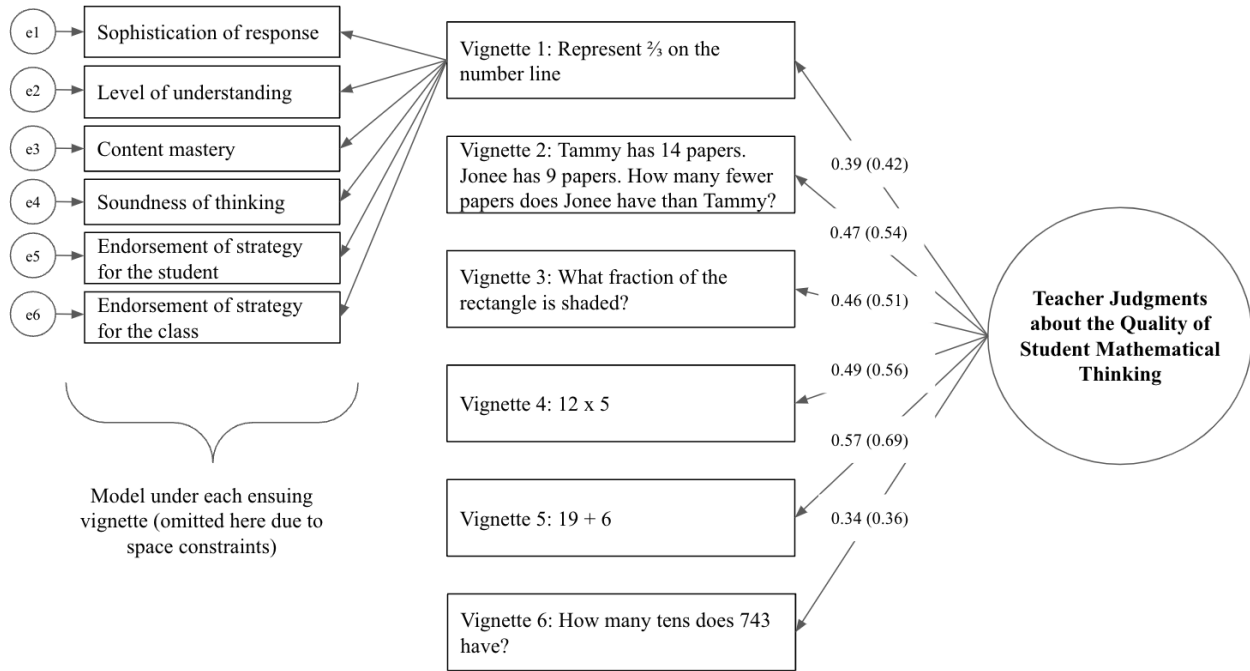
Table A1.1 Continued

Supporting Student Sensemaking (SSS)			
N/A	0 - No support of student sensemaking	1 - Support of student sensemaking	
No teacher utterance or teacher utterance is not responsive to prompt	Teacher does not invite student input or activity related to the original student statement. The teacher instead moves on to direct instruction, simple approval, or asks the class to move on to a <i>*different</i> problem or task. The nature of student input is such that students are asked to answer a question for which the teacher already has a clear and strict answer in mind. <i>* Different means the new task is unrelated/ unresponsive to where students might be at, e.g. “Let’s do another problem, class”)</i>	Teacher keeps the student idea in play somehow and/ or attends to student thinking in the classroom. The teacher “has an ear to the ground” and seeks student input, ideas, and reasoning OR assigns an additional task related to the original student idea or related to student thinking “on the ground”. The nature of student input is such that students are given the opportunity to engage in their own sense-making. E.g. “ <i>Who else agrees with this answer? Did anyone solve it a different way? Turn and talk to your partner about X’s answer.</i> ”	
Teacher Affirmation of Student Math Ideas (TA)			
N/A	0 - No teacher affirmation of student ideas/ work	1 - Generic teacher affirmation of student ideas/ work	2 - <i>Specific</i> teacher affirmation of student ideas/ work
No teacher utterance or teacher utterance is not responsive to prompt	Teacher provides <u>no praise</u> . Does not position students as competent and does not affirm what they did that was mathematically productive. Teacher does not affirm the student work or provide any type of praise.	Teacher assigns praise to the student in the form of positive content or tone, <u>but it is non-specific/ generic</u> (e.g. “ <i>Good job!</i> ” without specifying the mathematical idea or specific thing a student did). Teacher may only address the correctness of the answer or <u>student engagement</u> , but not the mathematical idea.	Describes the extent to which the teacher demonstrates the ability to position students as competent with affirmations <u>specifying what they did that was mathematically productive</u> . Teacher actively affirms the specific student work (e.g., calling out what is correct in the student work, noting a strength in student response).

Note: For any responses with N/A, teacher scores were estimated from their other vignettes scores. No teacher in the sample had more than two vignettes coded as N/A.

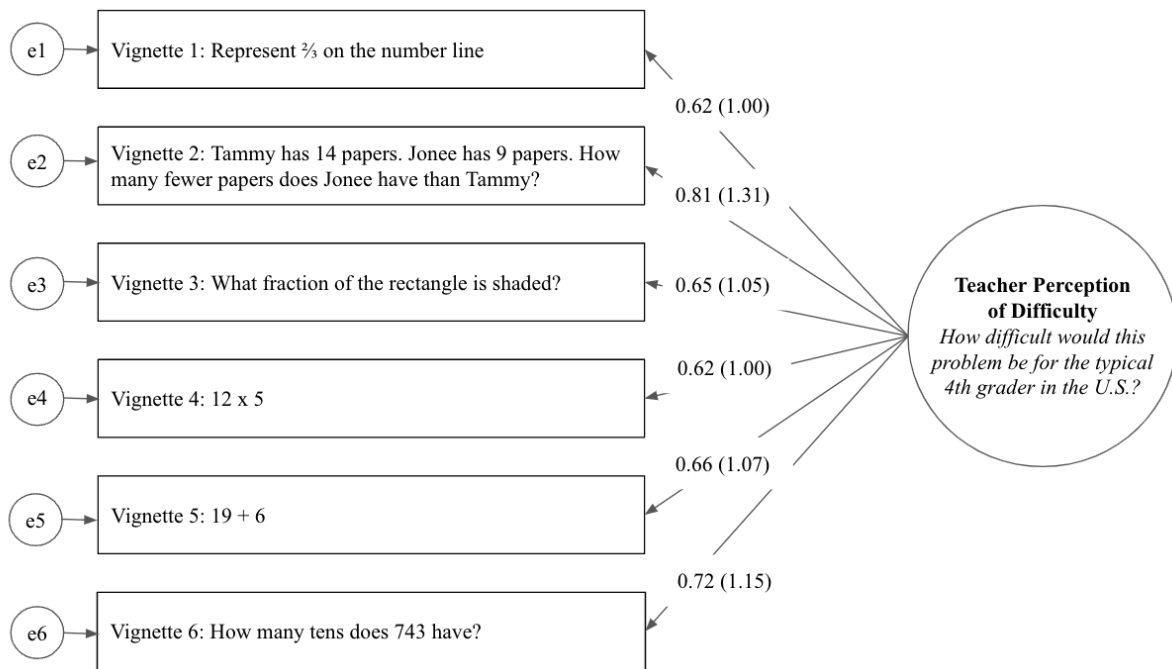
Appendix B.1

Figure B1.1 Conceptual Diagram and Confirmatory Factor Analysis for Measure of Teacher Judgments about the Quality of Student Mathematical Thinking



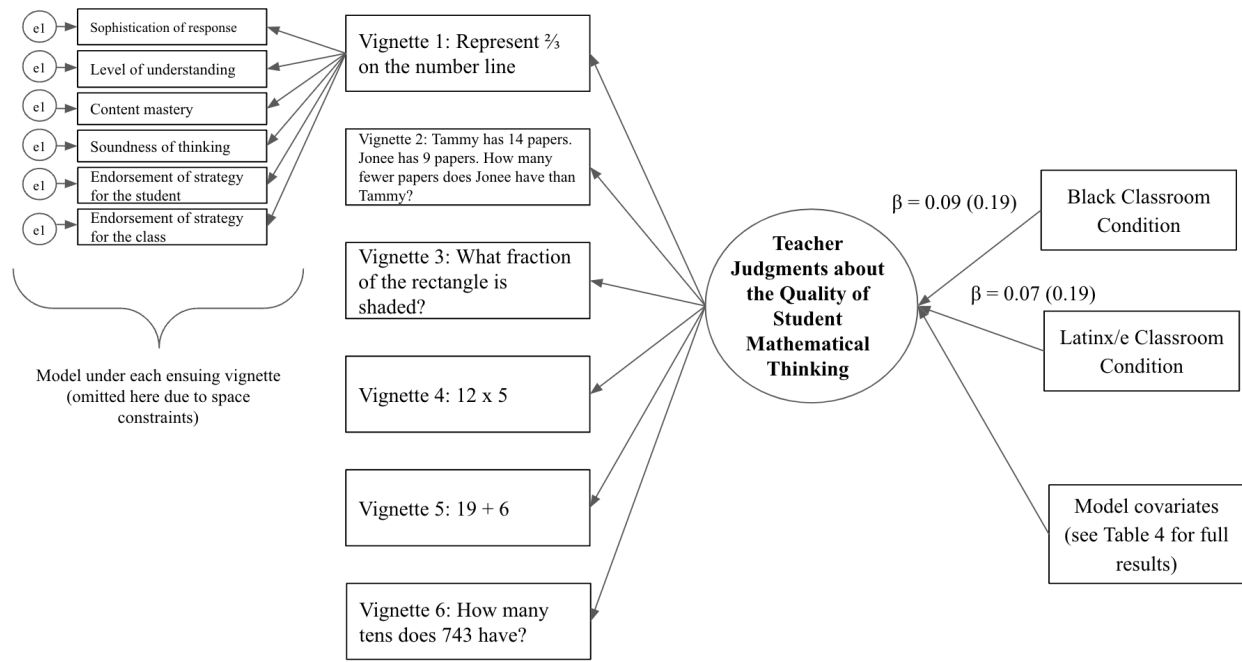
Note: Standardized coefficients (and unstandardized in parenthesis) are presented for the vignette level. $N=312$. Variances of measurement errors not presented. Covariances between correlated measurement error terms are not presented. Maximum likelihood (ML) estimator. Model fit indices: $\chi^2(580) = 1311.44, p < 0.00$; CFI = 0.925; TLI = 0.919; RMSEA = 0.06; SRMR = 0.07.

Figure B1.2. Conceptual Diagram and Confirmatory Factor Analysis for Measure of Teacher Expectations of Task Difficulty for the Student



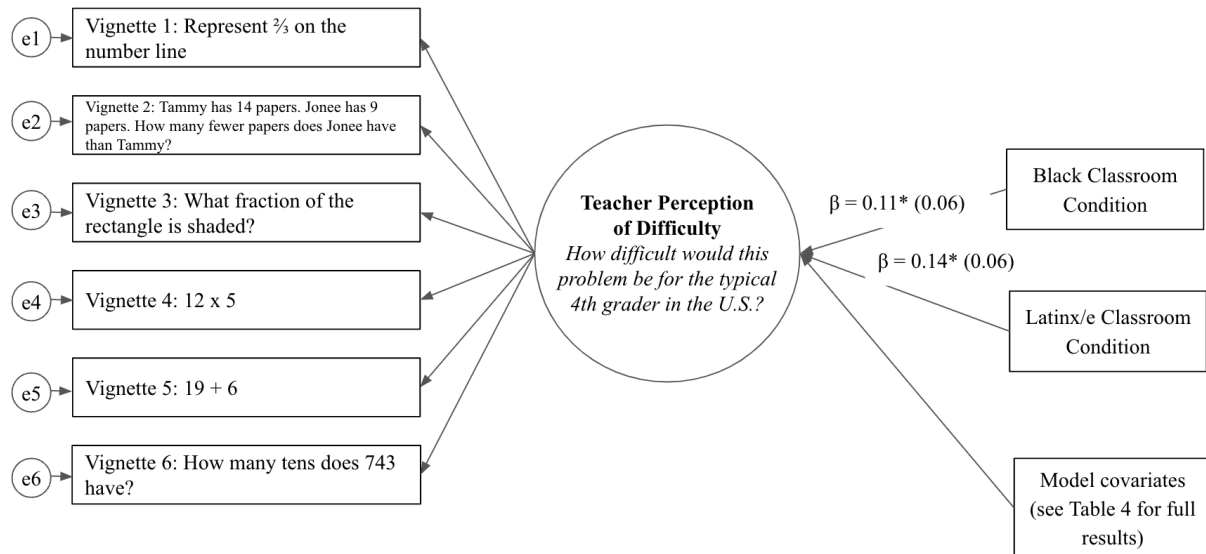
Note: Standardized coefficients (and unstandardized in parenthesis) are presented. $N=312$. Variances of measurement errors not presented. Diagonally weighted least squares (DWLS) estimator. Model fit indices: $\chi^2(9) = 17.085$, $p = 0.05$; CFI = 0.995; TLI = 0.991; RMSEA = 0.05; SRMR = 0.04.

Figure B1.3. Analytic Diagram of Structural Equation Model (SEM) for Relationship Between Black Classroom Condition and Latinx(e) Classroom Condition (Relative to White Classroom Condition) and Teacher Judgments about the Quality of Student Mathematical Thinking



Note: Full results showing coefficients for all covariates are shown in Table 1.4. $N=307$. Variances of exogenous variables, disturbance terms of endogenous variables, covariances between exogenous variables, and covariances between correlated measurement error terms not shown. Maximum likelihood (ML) estimator. Model fit indices: $\chi^2(1070) = 1912.88, p < 0.00$, CFI=0.913, TLI=0.908, RMSEA=0.05, SRMR = 0.06

Figure B1.4. Analytic Diagram of Structural Equation Model (SEM) for Relationship Between Black Classroom Condition and Latinx(e) Classroom Condition (Relative to White Classroom Condition) and Teacher Perception of Difficulty



Note: Full results showing coefficients for all covariates are shown in Table 1.4. $N=307$. Variances of exogenous variables, disturbance terms of endogenous variables, and covariances between exogenous variables not shown. Diagonally weighted least squares (DWLS) estimator. Model fit indices: $\chi^2(79) = 74.70$, $p=0.62$, CFI=1.00, TLI=1.00, RMSEA=0.00, SRMR = 0.04

Chapter 2

**Teachers' Mathematics Task Support and Attributions for Students' Struggles:
Investigating the Role of Student Race-Ethnicity, Teacher Background, and School Context**

Abstract

Teachers' diagnoses or explanations for student struggles relate to their practice in mathematics teaching, especially in classrooms serving racially minoritized learners. However, most of the studies on teachers' causal attributions for student struggles examine them along a success-failure binary rather than learning in progress, and little is known about relations between attributions, preservation of task cognitive demand, and teacher and school context factors. In this pre-registered study, I use a between-subjects design to randomize teachers to view an instructional vignette featuring the work of a hypothetical Black, Latinx/e, or White student stuck on a high-level math task. I find that student racial-ethnic condition showed no main effect on teachers' causal attributions, and educators who are women and teaching more advanced grade levels were less likely to endorse internal causal attributions for the student struggle. Further, teachers' instructional preferences to routinize, remediate, and extend the math task did not differ by the race-ethnicity of the student in the vignette but did differ descriptively according to the level of racial-ethnic segregation in the teaching context. Finally, teacher endorsement of statements attributing the student struggle to internal factors—but not external factors—positively related to their preferences to routinize, remediate, and extend the task, and these attribution-instruction relationships did not differ by racial-ethnic condition. The findings have implications for task-based mathematics instruction and research on teacher decision making in the face of student challenges.

Teachers' Mathematics Task Support and Attributions for Students' Struggles: Investigating the Role of Student Race-Ethnicity, Teacher Background, and School Context

When students are engaged in cognitively challenging task-based instruction, they are provided with opportunities to share their mathematical reasoning and develop conceptual understandings, fluency, and agency in mathematics (Bobis, 2024; Schoenfeld et al., 2023; Smith & Stein, 1998; Sullivan & Davidson, 2014). High-level cognitive demand characterizes math instruction where students are engaged and challenged to make purposeful connections to meaning, and student challenges or “impasses” have been described as necessary and even desirable components of the mathematics learning process that support cognitive engagement (Hiebert & Grouws, 2007; Schoenfeld et al., 2023; Stein et al., 2009; Warshauer, 2015).

Sustaining productive student engagement in cognitively demanding problem-solving tasks, particularly when students struggle, is challenging for teachers and complex to carry out (Sullivan et al., 2014), often resulting in the devolution of cognitive demand during task implementation both in terms of the coherence of the mathematics and student accountability for high-level processes and products (Stein & Smith, 1998). Problem routinization and remediation are common ways cognitive demand can decline during implementation, for example, when problem-worthy aspects of the task are removed by the teacher or when the focus of the task shifts to related but remedial mathematics (Estrella et al., 2020; Henningsen & Stein, 1997). Alternatively, after resolving a misconception, teachers can maintain or even increase cognitive demand by extending the task, for example, by encouraging students to refine and generalize their solutions (Bobis, 2024; Boston, 2012; Sullivan et al., 2006).

Whether teachers elect to routinize, remediate, or extend the cognitive demand of a task during implementation can depend on how they attribute or explain student difficulties. Teachers

must interpret student difficulties with and misconceptions about mathematics, explain or attribute their causes, and then develop a strategy for addressing them (Jackson, 2017; Jacobson et al., 2022; Wang & Hall, 2018; Wilhelm et al., 2017). Moreover, this attribution-instruction relationship may be more pronounced in classrooms with a greater composition of historically marginalized learners, likely from teachers assuming a greater responsibility for supporting students' learning (Wilhelm et al., 2017). Studies on the relationship between attributions and instructional preferences are limited, particularly during math task implementation when cognitive demand has great potential to vary as a function of teacher decisions (Warshauer, 2015).

Importantly, sociodemographic factors like student race-ethnicity, teacher, and school-level factors can shape both teachers' attributions and the cognitive demand of mathematics instruction, though systematic descriptive work is limited. Teachers' attributions for student struggles in mathematics can range from productive to unproductive explanations and differ according to salient identity factors such as student race-ethnicity (Copur-Gencturk et al., 2024; Froehlich et al., 2016; Jackson et al., 2017; Jacobson et al., 2022; Pollard, 2022). However, studies examining teachers' explanations or attributions for student difficulties often focus on success or failure in mathematics more broadly, rather than the ambiguous areas in between, such as student struggles while solving problems (Wang & Hall, 2018). When it comes to how sociodemographic factors shape the cognitive demand of mathematics instruction, prior experimental work has found teachers may be more likely to select remedial math tasks for Black students as compared to White students during the data use cycle (Pollard, 2022). The present study builds on this work by examining this phenomenon during math task implementation.

In this paper, I examine the relationship between teachers' causal attributions for a student struggling during task implementation and teacher preferences for instructional moves that alter

cognitive demand. Specifically, I investigate how teacher preferences for task remediation, routinization, and extension relate to their internal and external causal attributions for the student's struggle and the extent to which this attribution- instruction relationship might differ by student race-ethnicity. I also examine how K-8 teachers' attributions and preferences to remediate, routinize, and extend a math task each relate to student race-ethnicity, teacher characteristics, and school context factors. I do so via a survey presenting eight vignettes of common student "stuck points" during the implementation phase of mathematics tasks. I define a stuck point as a moment when a student has made some initial progress towards solving the problem of the task using a mathematically viable strategy but has reached an impasse due to a common math error or misconception (Borasi, 1994; VanLehn et al., 2003; Warshauer, 2015). Findings from this study contribute to a better understanding of the factors shaping cognitive demand in K-8 mathematics teaching, with implications for teacher education and development.

Theoretical Framework

Attribution theory provides a theoretical lens through which we can better understand how teachers explain student struggles in mathematics and justify ensuing instructional responses. Attributions are the perceived causes people use to explain achievement-related outcomes for themselves or others, which I focus on here. These explanations contain three dimensions: stability, controllability, and locus of causality (Weiner 1985; 2000; 2010). Stability refers to how persistent or fluctuating a factor is over time, controllability to the extent to which an individual has power or agency to shape the factor, and locus to the location of the factor relative to the individual, i.e. internal or external (Weiner, 1985; 2000; 2010). For example, a teacher might attribute a student's struggle in mathematics to their resources at home (mostly stable, uncontrollable, external), their

effort (unstable, controllable, internal), or to their innate mathematics ability (stable, uncontrollable, internal).

Importantly, the locus of causality may be prone to biases. The fundamental attribution error, sometimes referred to as correspondence bias, is a cognitive attribution bias that occurs when a person overemphasizes an internal locus compared to an external one (Gilbert and Malone, 1995). Attribution error and correspondence bias tend to be pronounced in achievement domains such as education (Reyna, 2008). Moreover, due to the salience of student identity factors such as race-ethnicity, the locus of causality may be a site sensitive to biases in mathematics teaching. For example, racialized discourses around a hierarchy of math achievement inherent to the student (internal, e.g. Martin, 2009) and perceived community deficits (external, e.g. Jackson et al., 2017) may make racial-ethnic biases more pronounced in the locus dimension of teachers' causal attributions.

Racial-ethnic biases trace back to stereotypes or cultural beliefs about the behavior and traits of social groups (Eagly & Chaiken, 1993; Hilton & von Hippel, 1996; Tajfel, 1970). A core function of stereotypes is group categorization, which is used to save time and cognitive resources when processing information and making sense of one's complex social world (Schneider, 2004). As a result of stereotyped thinking, members of a particular social group tend to be perceived in accordance with the stereotypes about their social group(s). Moreover, stereotyped associations may become more pronounced and "fill in the blanks" in ambiguous situations with low information (Dovidio, 2001). The effect of stereotypes on one's cognitive processing is biased judgments and behaviors, which are not based on the actual competencies and characteristics of a person, but rather on their group membership.

Review of the Literature

In this section, I provide an overview of the literature on teachers' causal attributions, instruction during math task implementation, and the attribution-instruction relationship, with attention to how they may be shaped by student race-ethnicity and teacher- and school-level factors. I conclude with an outline of how the present study extends existing work in this field.

Teachers' Causal Attributions

Teachers' attributions or explanations for students' successes or failures inform their expectations for learning and eventual student achievement (Dobbs & Arnold 2009; Graham 1991; Middleton & Spanias, 1999; Wilson et al., 2014). In the context of ambitious, reform-oriented mathematics teaching (Lampert et al., 2013), prior scholarship has described two sets of “frames” for how teachers explain or attribute the problem of differential student success, with relations to teacher- and school-level factors and ensuing instruction. Jackson et al., (2017) outlined productive and unproductive diagnostic frames teachers drew upon to explain student difficulties with math. When they take a productive diagnostic frame, teachers attribute student difficulties to learning opportunities, whereas they attribute difficulties to perceived student deficits, such as their innate ability or families, when they take an unproductive diagnostic frame. Jackson and colleagues (2017) found nearly 1 in 3 teachers offered entirely unproductive explanations, attributing students' struggles to student, family, and community deficits rather than to the opportunities to learn afforded by their prior instructional experiences. Understanding the locus of teachers' causal attributions beyond their productiveness would help to illustrate where they trace the sources of student challenges in mathematics, with implications for interventions.

The Influence of Student, Teacher, and School Factors on Teachers' Causal Attributions

Moreover, teachers' attributions may be informed by racial and ethnic stereotypes; as such, student race-ethnicity informs whether teachers attribute student performance to internal or external factors, both in general (e.g. Froehlich et al, 2016) and in the domain of mathematics (Copur-Gencturk et al., 2024; Jacobson et al., 2022; Pollard, 2022). When it comes to the math achievement for students of color, teachers' attributions may be informed by harmful stereotypes, such as the belief that Black and Latinx/e students are uniformly lazy and underprivileged (Reyna, 2000). In one experimental study, teachers were more likely to attribute errors to the student than to prior instruction when the student was Black (Pollard, 2022). Other experimental work has also demonstrated the tendency for teachers to more readily attribute student performance to internal factors rather than external ones for racially minoritized students. In a study of 400 elementary and middle school math teachers, teachers attributed student math performance to internal factors, such as their cognitive abilities, when they were told that Black and Hispanic students outperformed Asian and White students (Copur-Gencturk et al., 2024). But when they were told the opposite—that White and Asian students outperformed their Black and Hispanic peers—teachers tended to attribute to external sources such as parents and cultural differences that value math achievement. In sum, teachers may be more likely to attribute student mathematics achievement to internal factors for racially minoritized students, informed by racial-ethnic stereotypes.

Teacher-level factors also inform the nature of their attributions for student academic failure or success. Experienced teachers are more likely to link children's achievement to biological and stable factors that are outside a student's control, such as innate intellectual ability and student gender, as well as to external factors like family background, whereas novice teachers place a greater emphasis on the responsibility of the teacher and the student's effort (Georgiou, 2008). However,

novice teachers have also been found to be more likely than experienced educators to explain student failure with external factors, such as by blaming families or parents (Pirrone, 2012). When it comes to the role of a teacher's level of instruction (i.e. grade level taught), findings are inconsistent (Wang & Hall, 2018). Finally, teacher gender may inform their causal attributions, primarily due to their own experiences with attributional gender bias and the stereotype that boys are better at math (e.g. Espinoza et al., 2014). Taken together, these findings demonstrate how teacher background shapes their attributions more generally, and the present study builds on this work in the specific domain of mathematics teaching.

Teachers' Instructional Decisions While Implementing Demanding Tasks

A core instructional area in ambitious mathematics teaching involves launching and implementing math tasks with high-level cognitive demands involving meaningful connections (i.e. linking concepts, procedures, and their underlying meaning) and holding students accountable for *doing* mathematics (i.e. high-level mathematical processes and products) (Boston, 2012; Hiebert & Grouws, 2007; Schoenfeld et al., 2023; Stein & Lane, 1996; Stein et al., 1996). In their framework for describing the factors that shape the cognitive demand during the implementation phase of high-level mathematics tasks, Stein and Smith (1998) observe that teachers can devolve the cognitive demand of tasks in two ways: by routinizing the problem, and by pivoting to remediation. An example of how problem-worthy aspects of the task can become routinized is when the teacher provides a set of steps for the student to follow, thereby changing the “ask” and lowering the cognitive demand of the problem, e.g. “The problem is really asking us find x for this value of y , so we can plug in 6 for y and solve this equation here.” Another way the cognitive demand can be reduced is through remediation, which occurs when the focus of a math task shifts to a below-grade-level standard, such as identifying fractions on a labeled number line rather than having the

student make sense of and reason about part-whole relationships to label it themselves. In both cases, meaning-oriented mathematical connections of the task are pared down (with the latter even changing the problem to remedial math), and the student is no longer accountable for high-level processes or products.

On the other hand, extension of the task maintains or increases the cognitive demand of instruction through supplementation. For example, the teacher can extend the task by increasing the complexity of numbers or the operation or by encouraging refinements to and generalizations of student solutions, e.g. through extending prompts (Bobis, 2024; Boston, 2012; Sullivan et al., 2006). After addressing a student stuck point, a teacher may move to incrementally extend the task, rather than routinize or remediate, in order to build the student's understanding and help them progress towards mastery. For example, after supporting a student to locate the number 1 on a number line with "0" and " $\frac{1}{4}$ " labeled, the teacher may move to having the student locate the number 1 again, this time with only "0" and " $\frac{5}{3}$ " represented on the number line, extending the students' opportunities to reason with fractional values to support a measurement understanding of fractions.

The Influence of Student, Teacher, and School Factors on Math Teaching

When it comes to the role of student race-ethnicity in teachers' mathematics instructional preferences, recent experimental studies, while limited, generally show that teachers tend to hold biases against Black students compared with White students when responding to student work containing mathematical errors. In a randomized study with 90 grade 4–8 general education math teachers, Pollard (2022) found that teachers were more likely to prefer lower-quality curricular materials, remedial instruction, and student-centered instructional moves for Black students than for White students when responding to the student's error, while their evaluation of the mathematical work and feedback were equitable across student race. Another experiment involving student work

containing mathematical errors found that teachers were more likely to recommend Black students for special education compared to White students, and this finding did not extend to correct student mathematical work (Copur-Gencturk et al., 2022). Together, these studies highlight how racial biases shape teachers' curricular and pedagogical decisions, particularly when presented with student difficulties in mathematics. I build on this work by examining the potential for racial biases to shape the cognitive demand in teacher responses to student stuck points during task-based math instruction as well as by including Latinx/e students in addition to Black and White students.

Existing studies on the implementation of task-based mathematics instruction either do not systematically measure teacher and school background factors, or feature sample sizes too small for confident analyses, despite the fact that factors like teacher knowledge, aggregate student math achievement, level of instruction, and teaching experience relate to the cognitive demand of instruction and may influence how they respond to student stuck points (Gottfried et al., 2023; Hill et al., 2007).

The Relationship Between Teachers' Causal Attributions and Instructional Preferences

Teachers' attributions for students' academic struggles have been shown to relate to their instructional preferences and decisions (Reyna & Weiner, 2001; Wang & Hall, 2018). Specifically, teachers' instructional interventions vary depending on the extent to which they believe the student or teacher is responsible for the academic failure. Empirical and theoretical work has shown that when teachers believe students are in control of their academic failures, they tend to respond with punitive or retributive approaches (e.g. punishment), whereas when they believe the student is not in control of their academic failure (e.g. due to the student's innate abilities or prior instructional experiences), they tend to prefer utilitarian approaches, such as remediating instruction (Reyna & Weiner, 2001).

Studies specific to the context of ambitious mathematics teaching—which emphasizes student activation and engagement—show mixed findings for the relationship between teachers’ attributions for student struggles in math and their ensuing instructional responses. In a qualitative analysis of semi structured interviews with teachers, Jackson and colleagues (2017) examined the relationship between teachers’ attributions for student difficulties in mathematics and the rigor of ensuing instructional approaches, which were characterized as productive when the teacher reported supporting student engagement in cognitively demanding activity to build mastery, or unproductive when the mathematical rigor was reduced. The authors found teachers overwhelmingly reported lowering the cognitive demand of mathematics instruction when faced with student difficulties, and even when they explained student difficulties in terms of the instructional opportunities afforded to them, teachers inconsistently reported adjusting their instruction accordingly (Jackson et al., 2017). In other words, some teachers took responsibility for students’ misunderstandings and designed instruction to target them, while others did not.

In a related study going beyond teacher reflections and self-reports, Wilhelm and colleagues (2017) found that in classrooms with teachers who explained student difficulty in mathematics as a function of their instructional experiences, students were given greater opportunities to share their mathematical reasoning (also termed “providing”) compared to classrooms with teachers who explained difficulties as stemming from deficits believed to be inherent to students and their communities. Taken together, these two studies show that when teachers explain student struggles in terms of the affordances provided by their prior instructional experiences, they inconsistently adjust their practice accordingly. The present study builds on this literature on the attribution-action relationship in ambitious math teaching by homing in on the problem-specific, finer-grain interactional spaces of instruction.

The Influence of Student, Teacher, and School Factors on the Attribution-Instruction Relationship

Importantly, teachers' attributionally guided instructional preferences may be shaped by salient student demographics, such as race-ethnicity, as well as teacher- and school-level factors. In the context of ambitious math teaching, Wilhelm and colleagues (2017) found teachers' explanations or attributions for student difficulties were "more likely to be reflected in (or reflective of) their instructional practice" (p. 363) in classrooms with students of color and those with limited English proficiency status, but not student math achievement nor poverty level. This shows the attribution-action relationship may be stronger in classrooms with a greater composition of historically marginalized learners, as teachers may assume a greater responsibility for supporting students' learning. Teacher and school-level factors also predict both teachers' causal attributions and instructional preferences, the attribution-instruction link, and biases. For example, mean student math achievement moderates the relationship between teachers' attributions for student difficulty in math and student opportunities to share their reasoning in class, with this attribution-instruction link all but disappearing in classrooms with high proficiency. In terms of context factors, while we know interracial contact across groups can reduce implicit biases (Pettigrew & Tropp, 2006), we know little about the role of school integration as it relates to teachers' causal attributions, mathematics instructional practice, the relationship between the two, and biases therein.

The Present Study

I build on the existing literature in four key ways. First, the majority of scholarship on teachers' attributions or explanations for student performance in mathematics focuses on the success-failure binary rather than the ambiguous spaces in between (Wang & Hall, 2018). The field lacks any examination of teachers' attributional biases when addressing student stuck points during learning in progress, such as student work time during math task implementation. Second, while the

field has qualitative and descriptive evidence describing the relationship(s) between teachers' attributions and instructional approaches in the context of ambitious math teaching more generally (Jackson et al., 2017) and with respect to specific classroom discourse practices (Wilhelm et al., 2017), more work is needed to understand this relationship during task implementation. Third, I extend prior work examining biases in teachers' instruction in the context of ambiguous student work to include Latinx/e students in addition to Black and White students, as well as comprehensive teacher- and school-level factors, such as segregation of the school context. Finally, the most relevant studies examining teacher attributions for student difficulties in the context of ambitious math teaching have characterized attributions in terms of their productiveness rather than according to the locus of causality (e.g. internal or external). The present study incorporates a longer research tradition with more established measures from attribution theory, with implications for where the responsibility for remediation rests, e.g. the student, parents/ caregivers, teachers. Thus, in the present study, I ask:

1. How do K-8 teachers' internal and external attributions relate to student race-ethnicity, teacher characteristics, and school context factors?
2. How does teacher support for remediation, routinization, and extension activities relate to student race-ethnicity, teacher characteristics (e.g. mathematical knowledge for teaching, MKT) and school context factors?
3. How does teacher support for remediation, routinization, and extension activities in their next instructional move relate to their internal and external causal attributions for the student's struggle? Do these relationships differ by student race-ethnicity?

Methods

Study Design and Procedures

Teachers responded to an online survey administered via Qualtrics in the spring of 2025. Participants were screened using two questions asking whether they have been a certified teacher of record in the U.S. with responsibility for instruction in a classroom that is between grades K and 8, and whether they teach mathematics as part of their teaching responsibilities. The survey informed participants they would be responding to teaching scenarios or “vignettes” that take place during student work time on math problems, and that the scenarios featured the same 4th grade student and their work. Although we sacrificed the face validity associated with classroom observation data with our use of vignettes, this approach enabled us to make causal inferences about the role student race-ethnicity may play in teachers’ attributions and their ensuing instructional preferences.

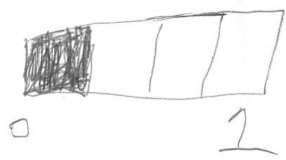
At the start of the survey, participants were randomly assigned to one of three conditions using a blocked survey design. This study was pre-registered in the Open Science Framework (OSF)⁵ and employed a between-subjects design in which participants were randomly assigned to view instructional vignettes featuring the work of a hypothetical Black (condition A, $n = 85$ teachers), White (condition B, $n = 81$ teachers), or Latina girl (condition C, $n = 79$ teachers). Randomly assigning teachers to experience vignettes in only one condition rather than across conditions (a between- rather than within- subject design) diminished the potential for social desirability bias or participants becoming aware that the study was about race (Hofmann et al., 2005). Vignettes were identical save for the race-ethnicity of the student as operationalized by student name. Names were selected using those commonly associated with particular racial-ethnic groups, but not socioeconomic status (Gaddis, 2017a, b). Teachers were told to assume that the student is a typical

⁵ OSF pre-registration link: https://osf.io/q2yz8/overview?view_only=9c2ca47a337a4e739a313b9319824d8e

student in a general education classroom, is from a family at the U.S. median household income level that does not qualify for free or reduced price lunch, has no special needs nor cognitive disabilities, is fluent in English and not considered an English language learner, and performs at the average mathematics proficiency level in the U.S. based on standardized testing. The purpose of this statement was to reduce the potential for teacher assumptions about other factors that may contribute to student stuck points across conditions. All teachers retained in the final sample passed the survey's attention check and commitment check questions.

The survey presented teachers with eight instructional vignettes featuring a mathematics task and the work of a student who has reached a stuck point, defined for participants simply as a place where the student has stopped making progress on the task. Importantly, I define a stuck point as a moment during task implementation where the student has made some initial progress towards solving the problem of the task using a mathematically viable strategy but has made a mathematical error, conceived broadly and comprehensively following Borasi (1994), leading to an impasse (VanLehn et al., 2003; Warshawer, 2015). Tasks were created using free online resources by Illustrative Mathematics ® and the author's own task creations in Adobe Photoshop ©. Additionally, tasks were designed with input from a focus group of math educators and former teachers. The design features of the math tasks and accompanying prompts featured in the vignettes were further developed and finalized after holding cognitive interviews with five current or former K-8 elementary school teachers.

Figure 2.1. Sample vignette with task, student work, and prompt

Prompt	Mathematics Task and Student Work
Your class is working on the task below about identifying fractional parts on a number line. You notice a student has drawn a fraction diagram with four equal parts and $\frac{1}{4}$ shaded in out of 1 whole, but they have not labeled the number line, as shown on their paper below.	[Condition A] Name <i>Tanisha Washington</i> Locate 1 on the number line. Label the point. Be as exact as possible  
What would you say to this student to support them in locating 1 on the number line?	

Note: Student first and last names were used to operationalize race-ethnicity across the three experimental conditions, and they were selected from a list of names commonly associated with particular racial-ethnic groups (Gaddis, 2017a, b). The vignette featured here shows the name for the condition showcasing the work of a hypothetical Black student. In the condition featuring a hypothetical White student, race-ethnicity was operationalized using the name “Katie Walsh.” In the Latinx/e condition, the name was “Daniela Velasquez.”

Vignettes were standardized to describe the topic students in class were working on, the math task with instructions and the student’s name at the top, the student’s work in progress, and a description of what they had completed so far as well as their stuck point. See Table A2.1 in Appendix A2 for the content, prompts, and features of each vignette. An example of one of the vignettes as seen by teachers is shown in Figure 2.1. In all vignettes, teachers were asked to respond verbally to the question, “*What would you say to this student to support them in [goal of the task]?*” and to record themselves speaking out loud. Teacher responses were recorded and transcribed using Voiceform software embedded into Qualtrics. Following each vignette were three Likert questions asking teachers about the extent to which they would remediate, routinize, or extend the task based on the student’s response. At the end of the survey, teachers responded to background questions

and nine Likert questions about internal and external attributions for the student's struggles in mathematics.

Sample

The final sample included $N=245$ teachers. A power analysis ($\alpha = .05$, power = .80) conducted using PowerUp indicated that a sample of $N=245$ participants would yield a minimum detectable effect size (MDES) of 0.27 for the interaction models.

Table 2.1. Sample Characteristics for Full Sample and Across Conditions

	Full Sample $N=245$	Condition A $n=85$	Condition B $n=81$	Condition C $n=79$	Pearson Chi-square or ANOVA
	N (%)	n (%)	n (%)	n (%)	
Gender					
Woman	137 (55.9%)	53 (62.4%)	42 (51.9%)	42 (53.2%)	$\chi^2(2) = 2.41$, $p = 0.33$
Man	101 (41.2%)	30 (35.3%)	38 (46.9%)	33 (41.8%)	$\chi^2(2) = 2.33$, $p = 0.31$
Nonbinary	7 (2.9%)	2 (2.4%)	1 (1.2%)	4 (5.1%)	$\chi^2(2) = 2.23$, $p = 0.31$
Years of Experience	10.3 (8.1)	11.0 (8.4)	10.4 (8.3)	9.4 (7.5)	$\chi^2(2) = 2.25$, $p = 0.32$
Grade level taught					$\chi^2(18) = 14.75$, $p = 0.70$
Kindergarten	8 (3.3%)	3 (3.5%)	2 (2.5%)	3 (3.8%)	
1 st grade	16 (6.5%)	6 (7.1%)	6 (7.4%)	4 (5.1%)	
2 nd grade	25 (10.2%)	7 (8.2%)	6 (7.4%)	12 (15.2%)	
3 rd grade	38 (15.5%)	17 (20.0%)	10 (12.3%)	11 (13.9%)	
4 th grade	47 (19.2%)	14 (16.5%)	16 (19.8%)	17 (21.5%)	

Table 2.1. Continued

	Full Sample N=245	Condition A n=85	Condition B n=81	Condition C n=79	Pearson Chi- square or ANOVA
	N (%)	n (%)	n (%)	n (%)	
5 th grade	49 (20.0%)	14 (16.5%)	23 (28.4%)	12 (15.2%)	
6 th grade	24 (9.8%)	9 (10.6%)	6 (7.4%)	9 (11.4%)	
7 th grade	14 (5.7%)	5 (5.9%)	6 (7.4%)	3 (3.8%)	
8 th grade	23 (9.4%)	10 (11.8%)	6 (7.4%)	7 (8.9%)	
High School ⁶	1 (0.4%)	0 (0.0%)	0 (0.0%)	1 (1.3%)	
Teacher race- ethnicity					
Black	45 (18.4%)	16 (18.8%)	10 (12.3%)	19 (24.1%)	$\chi^2(2) = 3.67$, p = 0.16
Latinx/e	34 (13.9%)	6 (7.1%)	14 (17.3%)	14 (17.7%)	$\chi^2(2) = 5.01$, p = 0.08
Native Am.	2 (0.8%)	1 (1.2%)	0 (0.0%)	1 (1.3%)	$\chi^2(2) = 1.00$, p = 0.77
AAPI	6 (2.4%)	3 (3.5%)	1 (1.2%)	2 (2.5%)	$\chi^2(2) = 0.92$, p = 0.78
White	148 (60.4%)	55 (64.7%)	54 (66.7%)	39 (49.4%)	$\chi^2(2) = 6.01$, p = 0.05
Multiracial	10 (4.1%)	4 (4.7%)	2 (2.5%)	4 (5.1%)	$\chi^2(2) = 0.82$, p = 0.72
MKT factor score	0 (1.00)	0.07 (1.0)	- 0.03 (1.0)	- 0.05 (1.0)	F(2, 242) = 0.37, p = 0.69

⁶ We retained $n=1$ high school teacher who met our screening criteria of having taught K-8 math in the past three years

Table 2.1. Continued

	Full Sample	Condition A	Condition B	Condition C	Pearson Chi-square or ANOVA
	N=245	n=85	n=81	n=79	
	N (%)	n (%)	n (%)	n (%)	
School Demographics (<i>n</i> =217)	Mean				
Metro Area SES	0.19	0.20	0.21	0.17	$F(2, 214) = 0.10$, $p = 0.90$
Math Achievement	0.00	-0.002	0.041	-0.023	$F(2, 214) = 1.76$, $p = 0.17$
Avg. W-B/L Segregation	0.34	0.36	0.33	0.34	$\chi^2(2) = 1.37$, $p = 0.50$

Note: Experimental conditions A, B, and C consisted of a hypothetical student who was a Black, White, or Latina girl, respectively. AAPI stands for Asian American and Pacific Islander. Column percentages are shown in parentheses. *Metro Area SES* is a composite score describing the socioeconomic status of school context metro area. *Math Achievement* is the cohort scale mean math achievement in the school context metro area. *Avg. W-B/L Segregation* represents the average of the school context metro area White-Black and White-Hispanic segregation indices relative to the local population, from 0 to 1, with 1 being total segregation. Means and SD's are provided for years of teaching experience. Means are provided for school demographic variables. We ran one-way ANOVA tests to compare means across conditions for school demographic variables except segregation. For average years of teaching experience and the segregation index, which had non-normal distributions, I ran Kruskal-Wallis tests, corrected for ties. For our categorical variables, we used Chi-Square tests and report the Fisher's exact test results when cells had few (<5) frequencies.

The majority of teachers sampled were women (56%), which is a lower proportion of compared to the national average in U.S. public schools (National Center for Education Statistics, NCES, 2023). Teachers reported an average of 10.3 years of teaching experience. The level of instruction for 75% of teachers was between grades 2-6. Most teachers identified as White (60%), Black (18%), or Latinx/e (14%), which shows our sample had a more racially and ethnically diverse distribution compared to U.S. public schools (NCES, 2023). Teacher scores on the Mathematical Knowledge for Teaching (MKT) assessment (Hill et al., 2008) were comparable to that of a national

sample collected in 2016 (Hill et al., 2019). In terms of the demographics of their school contexts, teachers taught in metro areas with above-average socioeconomic status, average mathematics achievement, and low to moderate levels of White-Black and White-Hispanic segregation.

To test for potential differences in teacher characteristics across our three conditions, I ran a series of Pearson Chi-Square and ANOVA tests and found that by chance, Latinx/e teachers may have been somewhat underrepresented in the condition featuring the work of a Black student, $\chi^2(2) = 5.01, p = 0.08$. White teachers may have also been underrepresented in the condition featuring the work of a Latinx/e student, $\chi^2(2) = 6.01, p = 0.05$. However, these differences did not reach the threshold of statistical significance, and alternative specifications did not substantively change the findings. There were no notable differences across conditions along other dimensions. See Table 2.1 for a summary of the teacher sample across conditions.

Measures and Scores

Positive Attributions

The set of questions capturing teachers' positive internal and external attributions for the student stuck point was informed by Jacobson et al.'s (2022) Attributions of Mathematical Excellence Scale (AMES) as well as Copur-Gencurk and colleagues' (2024) paper investigating teachers' attributional race and gender bias. Teachers responded to the prompt: "For the 4th grade student whose mathematical stuck points you considered in all of the questions in this survey, please indicate the extent to which you agree with the following statements." Teachers rated their responses on a scale of 1 (disagree strongly) to 5 (agree strongly). See Table 2.2 for a summary of the items measuring teachers' positive causal attributions as well as descriptors for each item's locus, stability, and control. Importantly, the focus of this study was on teachers' internal and external causal attributions to understand where they locate the source or locus of student difficulties rather

than on stability or control. Excluding the negatively worded items (#1 and #5) improved the internal coherence of the measure for teacher attributions internal to the student from $\alpha = 0.67$ to $\alpha = 0.77$. Similarly, for external attributions, excluding the negatively worded item (#6) improved the internal coherence of the measure from $\alpha = 0.61$ to $\alpha = 0.68$. I estimated factor scores for teachers' positive internal attributions using items 2-4 and positive external attributions using items 7-9.

Table 2.2. Items Measuring Teachers' Causal Attributions

Item	Statement	Stability and control
Positive attributions <u>internal</u> to the student		
1*	This student may not have put forth their best effort on these math tasks	controllable, unstable
2	This student likely paid close attention and followed directions in class.	controllable, unstable
3	This student is likely engaged and motivated when it comes to mathematics.	controllable, unstable
4	This student has a natural ability with mathematical reasoning.	uncontrollable, stable
5*	This student may not be ready for word problems or other challenging math tasks.	uncontrollable, stable
Positive attributions <u>external</u> to the student		
6*	This student likely did not get support with mathematics at home	uncontrollable, stable
7	This student likely enjoyed plentiful resources for learning mathematics outside of the classroom.	uncontrollable, stable
8	This student likely received strong prior instruction from their teacher on these mathematics topics.	uncontrollable, unstable
9	This student likely has experience solving word problems or other challenging math tasks in class	uncontrollable, unstable

Note: Asterisks indicate reverse-coded items

Teacher Preferences for Task Remediation, Routinization, and Extension

After viewing each vignette, teachers were given the following prompt “Based on this student’s work on this task, please indicate the extent to which you agree with the following statements,” followed by three statements asking them to rate their instructional preferences with

respect to remediating, routinizing, and extending the task on a scale of 1 (disagree strongly) to 5 (agree strongly). All items were tailored to the specifics of the mathematics of the task featured in each vignette. For example, the first vignette, which features a task that has a student locate the number 1 on a number line labeled with the numbers 0 and $\frac{1}{4}$, the remediation item measures teacher agreement to have the student identify fractional parts on a picture diagram (e.g. shaded circle) rather than a number line. For routinization, this item measures teacher agreement to label the number line for the student by counting by 4ths. For both of these remediation and routinization items, we interpret low agreement as the teacher wishing to maintain cognitive demand rather than diminish it. For task extension on the first vignette, we measure teacher agreement to have the student locate the number 1, this time with only the numbers 0 and $\frac{5}{3}$ labeled. See Tables A2.2-A2.4 in Appendix A2 for a summary of the items measuring teacher preferences for task remediation, routinization, and extension, respectively. The eight items measuring teacher preference for task remediation across the eight vignettes demonstrated adequate internal coherence ($\alpha = 0.77$), as did the items for routinization ($\alpha = 0.73$) and extension ($\alpha = 0.88$). Finally, a comparative model fit for nested models revealed that a 3-factor solution fit the data best compared to a 2-factor solution, supporting three distinct factors for remediation, routinization, and extension, $\Delta\chi^2(22) = 54.14, p < 0.001$ (Satorra, 2000; Kline, 2016). Thus, factor scores for teacher preference for task remediation, routinization, and extension were estimated for each teacher, respectively.

Teacher Background

I use an ordinal variable to measure years of teaching experience. For the level of instruction, I created an ordinal variable for grade from 0-9, with 0 being kindergarten, 1 being 1st grade and so on and so forth, ending with 8 for 8th grade and 9 for high school (for the one

secondary level teacher we retained, see above). Teacher gender was measured as woman, man, or non-binary and recoded as dummy variables for each. For teacher race-ethnicity, I used dummy variables to represent the following self-reported identities: Black, Latinx/e, White, Asian American or Pacific Islander, Native American, or multiracial. Due to small sample sizes for some of these identities, I could not conduct interaction analyses between condition and teacher race. These small sample sizes and a lack of theory also prevented me from interpreting the coefficients for teacher race-ethnicity in the results below.

Teacher Mathematical Knowledge for Teaching (MKT), which is the “complex mathematical understanding, skill, and fluency used in the work of helping others learn mathematics” (Thames & Ball, 2010, p. 228), has been shown to predict the rigor of the mathematical content present in instruction (Hill et al., 2007). As such, it is included as a covariate in the models so that we may measure teachers’ causal attributions and instructional preferences above and beyond their math knowledge for teaching. Teacher MKT was measured using 13 items selected from the Mathematical Knowledge for Teaching project developed by the Learning Mathematics for Teaching project at the University of Michigan (Hill, Schilling & Ball, 2008). Informed by item response theory, I selected items based on their informativeness and difficulty across different levels of teacher proficiency. The internal coherence across the 13 items was $\alpha = 0.61$, and I estimated a factor score for each teacher.

Teaching Context Covariates

Prior scholarship underscores the importance of disentangling biases from the social reality of racialized disparities in student achievement (Jussim & Harber, 2005). Teacher biases may be informed by students’ prior mathematics achievement, socioeconomic status, gender, student diagnostic labeling, and race-ethnicity (Peterson et al., 2016). As such, we asked teachers for their

school zip code and cross-referenced that with the Stanford Education Data Archive (SEDA) to obtain the following covariates from the metro area of their school contexts: a composite score describing the metro locale's socioeconomic status, a cohort scale mean of the math achievement in the metro area, and measures of Black-White and Hispanic-White school integration relative to the racial-ethnic composition of the locale (Fahle et al., 2025). To measure the degree of segregation or integration in the school context, we use the two-group information theory index, which assesses the degree to which schools' population diversity between two racial-ethnic groups deviates from the diversity of the local geographic area, with 0 signaling no segregation and 1 signaling total segregation (Owens et al., 2022). Because the indices for Black-White and Hispanic-White school integration are highly correlated ($r=0.77$), creating concerns about collinearity, I averaged both of these to create a single index to use in the models, interpreted as the degree of school segregation between White students and Black and Hispanic students, relative to the local population.

Analysis

For all research questions, I describe and interpret basic descriptives before sharing findings from the regression analyses specified below. In each model, student race-ethnicity is entered as a fixed effect and the reference group is the White condition. Additionally, in all models, **X** represents the set of covariates described above, including a dummy variable indicating the teacher passed a manipulation check and correctly noted the race-ethnicity of the student featured in the vignettes, and error describes a random error variance term.

RQ1. How do K-8 teachers' internal and external attributions relate to student race-ethnicity, teacher characteristics and school context factors?

For this question, I ran two separate multiple regression / ANCOVA models each with the dependent variable, Y_j , representing the factor score for teachers' positive internal and external attributions, respectively, as represented in equation 1 for the two models:

$$\textbf{Equation 1} \quad Y_j = B_{0j} + B_1\textit{Black}_j + B_2\textit{Latin}_j + B_3X_j + \textit{error}_j$$

Where j =participant and B_1 and B_2 are the coefficients measuring the main effects of bias in teachers' positive internal and external attributions, respectively.

RQ2. How does teacher support for remediation, routinization, and extension activities relate to student race-ethnicity, teacher characteristics and school context factors?

To answer this question, I ran three separate multiple regression / ANCOVA models with factor scores for teacher preference for task remediation, routinization, and extension each as the outcome, Y_j , respectively (equation 2):

$$\textbf{Equation 2} \quad Y_j = B_{0j} + B_1\textit{Black}_j + B_2\textit{Latin}_j + B_3X_j + \textit{error}_j$$

Where j =participant and B_1 and B_2 are the coefficients measuring the main effects of bias in teacher preference for task remediation, routinization, and extension, respectively.

RQ3. How does teacher support for remediation, routinization, and extension activities in their next instructional move relate to their internal and external causal attributions for the student's struggle? Do these relationships differ across the conditions?

I created a series of models interacting teachers' positive internal causal attributions and racial-ethnic condition (equations 3-6):

Equation 3
$$Y_j = B_{0j} + B_{1j}InternalAtt_j + B_{2j}Black_j + B_{3j}Latinx_j + B_{4j}InternalAttXBlack_j + B_{5j}InternalAttXLatinx_j + B_{3j}X_j + error_j$$

Where Y_j represents factor scores for teacher preference for task remediation, routinization, and extension, respectively, for each of the three separate models.

The relationship between teachers' positive internal causal attributions and task instructional preferences for each condition can be interpreted by the following equations:

Equation 4 White condition:
$$Y_j = B_{0j} + B_{1j}InternalAtt_j + B_{3j}X_j + error_j$$

Equation 5 Black condition:
$$Y_j = (B_{0j} + B_{2j}) + (B_{1j} + B_{4j})InternalAtt_j + B_{3j}X_j + error_j$$

Equation 6 Latinx condition:
$$Y_j = (B_{0j} + B_{3j}) + (B_{1j} + B_{5j})InternalAtt_j + B_{3j}X_j + error_j$$

Similarly, I created a series of models interacting teachers' positive external causal attributions and racial-ethnic condition (equations 7-10):

Equation 7
$$Y_j = B_{0j} + B_{1j}ExternalAtt_j + B_{2j}Black_j + B_{3j}Latinx_j + B_{4j}ExternalAttXBlack_j + B_{5j}ExternalAttXLatinx_j + B_{3j}X_j + error_j$$

Where Y_j represents factor scores for teacher preference for task remediation, routinization, and extension, respectively, for each of the three separate models.

The relationship between teachers' positive external causal attributions and task instructional preferences for each condition can be interpreted by the following equations:

Equation 8 White condition:
$$Y_j = B_{0j} + B_{1j}ExternalAtt_j + B_{3j}X_j + error_j$$

Equation 9 Black condition:
$$Y_j = (B_{0j} + B_{2j}) + (B_{1j} + B_{4j})ExternalAtt_j + B_{3j}X_j + error_j$$

Equation 10 Latinx condition:
$$Y_j = (B_{0j} + B_{3j}) + (B_{1j} + B_{5j})ExternalAtt_j + B_{3j}X_j + error_j$$

Descriptive analyses and regressions were completed using Stata version 17.0, and factor analysis was completed using RStudio version 2023.12.1+402 with the lavaan package for robust model fit indicators to accommodate polytomous response data.

Findings

Descriptive Results

On average, teachers showed moderately high agreement with statements that attributed the student's struggle in mathematics to positive internal factors, such as a natural mathematical ability, motivation, and engagement (mean= 3.31, SD= 0.91). Teachers were slightly less likely to agree with positive, externally focused reasons for student struggle. On average across conditions, teachers showed moderate agreement with statements that attributed the student's struggle in mathematics to external factors, such as experiences with strong prior math instruction and solving challenging math tasks and plentiful resources to support math learning at home (mean= 3.11, SD= 0.89). See Table 2.3 for a summary.

Table 2.3. Average teacher agreement with positive internal and external causal attributions

	Full Sample <i>N</i> =245	Condition A (Black) <i>n</i> =85	Condition B (White) <i>n</i> =81	Condition C (Latinx/e) <i>n</i> =79
	Mean (SD)			
Positive Internal Causal Attributions	3.31 (0.91)	3.30 (0.96)	3.24 (0.93)	3.40 (0.81)
Positive External Causal Attributions	3.11 (0.89)	3.12 (0.92)	3.05 (0.88)	3.17 (0.85)

Looking now to teachers' instructional preferences, on average across the eight vignettes, teachers showed relatively strong support for instructional approaches that routinized (mean= 4.15, SD= 0.57) and remediated (mean= 4.03, SD= 0.64) the task. They showed moderately strong

support for instructional approaches that extended the task (mean= 3.80, SD= 0.87). See Table 2.4 for a summary.

Table 2.4. Average teacher agreement with remediation, routinization across the eight vignettes

	Full Sample <i>N</i> =245	Condition A (Black) <i>n</i> =85	Condition B (White) <i>n</i> =81	Condition C (Latinx/e) <i>n</i> =79
	Mean (SD)			
Remediation	4.03 (0.64)	4.05 (0.58)	3.95 (0.67)	4.08 (0.68)
Routinization	4.15 (0.57)	4.19 (0.55)	4.15 (0.55)	4.11 (0.62)
Extension	3.80 (0.87)	3.80 (0.88)	3.80 (0.82)	3.81 (0.93)

RQ1: How do K-8 teachers' internal and external causal attributions relate to student race-ethnicity, teacher characteristics and school context factors?

Looking now at how the race-ethnicity of the student featured in the vignette and teacher and school factors individually predict teachers' positive causal attributions, we do not find statistically significant differences in teachers' internal nor external causal attributions by racial-ethnic condition. We find teachers who are women ($\beta = -0.30, p = 0.028$) and in more advanced levels of instruction ($\beta = -0.06, p = 0.081$) may be less likely to attribute the student struggle to positive factors internal to the student, all other factors constant. We did not find any other teacher and school context factors to relate to their causal attributions. See Table 2.5 for a summary.

Table 2.5. Teachers' Positive Internal and External Causal Attributions by Condition

	Internal Causal Attributions		External Causal Attributions	
Black Student	0.07 (0.16)	0.00 (0.16)	0.07 (0.16)	0.01 (0.17)
Latinx/e Student	0.17 (0.16)	-0.05 (0.16)	0.13 (0.16)	-0.02 (0.17)
Passed Manip. Check		0.16 (0.13)		-0.07 (0.14)
Grade Level Taught		-0.06 [†] (0.03)		-0.03 (0.03)
Years Teaching		0.00 (0.01)		0.01 (0.01)
Woman		-0.30* (0.13)		-0.11 (0.14)
Non-Binary		-0.12 (0.52)		-0.69 (0.55)
MKT Score		0.03 (0.07)		-0.08 (0.07)
Black Teacher		0.17 (0.17)		0.21 (0.18)
Latinx/e Teacher		0.26 (0.20)		0.25 (0.21)
AAPI Teacher		0.71 [†] (0.40)		0.35 (0.42)
Native Am. Teacher		0.45 (0.68)		0.26 (0.72)
Multiracial Teacher		0.20 (0.40)		0.37 (0.43)
Metro Area SES		-0.14 (0.16)		0.04 (0.16)
Math Achievement		-0.14 (0.40)		0.09 (0.42)
Avg. W-B/L Segregation		0.23 (0.48)		0.41 (0.50)
	R ² = 0.01	R ² = 0.08	R ² = 0.003	R ² = 0.05

*** $p < .001$, ** $p < 0.01$, * $p < 0.05$, $^{\dagger}p < 0.10$

Note: Coefficients are shown with standard errors in parentheses. $N=245$ for all unconditional models and $N=217$ for all conditional models due to observations missing covariate data from SEDA crosswalk. AAPI stands for Asian American and Pacific Islander. *Metro Area SES* is a composite score describing the socioeconomic status of school context metro area. *Math Achievement* is the cohort scale mean math achievement in the school context metro area. *Avg. W-B/L Segregation* represents the average of the school context metro area White-Black and White-Hispanic segregation indices relative to the local population, from 0 to 1, with 1 being total segregation. For attribution measures, we interpret coefficients as standardized differences in factor scores.

RQ2: How does teacher support for remediation, routinization, and extension activities relate to student race-ethnicity, teacher characteristics and school context factors?

Vignette student race-ethnicity did not relate to teacher agreement with instructional approaches that remediated, routinized, or extended the task; in other words, we did not observe racial-ethnic biases in the extent to which teachers preferred to alter the cognitive demand of the task. In terms of how teacher characteristics related to their instructional preferences, we found that teaching a more advanced grade level was positively associated with a preference for task extension ($\beta = 0.07, p = 0.025$). Looking at school context factors, the greater the average White-Black and White-Hispanic segregation relative to the local population, the more likely teachers agreed to extend the task ($\beta = 1.14, p = 0.021$). We did not find any other teacher and school context factors to relate to their preferences for task remediation, routinization, and extension. See Table 2.6 for a summary.

Table 2.6. Teachers' Likelihood of Remediation, Routinizing, and Extending by Condition

	Remediation		Routinization		Extension	
Black Student	0.15 (0.16)	0.19 (0.18)	0.07 (0.16)	0.13 (0.18)	0.00 (0.16)	0.09 (0.17)
Latinx/e Student	0.20 (0.16)	0.24 (0.18)	-0.07 (0.16)	0.02 (0.18)	0.00 (0.16)	0.14 (0.17)
Passed Manip. Check		-0.13 (0.15)		-0.06 (0.14)		0.03 (0.14)
Grade Level Taught		0.01 (0.04)		0.02 (0.03)		0.07* (0.03)
Years Teaching		0.00 (0.01)		0.00 (0.01)		0.00 (0.01)
Woman		0.07 (0.15)		0.04 (0.15)		0.00 (0.14)
Non-Binary		0.75 (0.58)		0.29 (0.57)		0.58 (0.53)

Table 2.6. Continued

	Remediation		Routinization		Extension	
MKT Score		-0.01 (0.07)		0.04 (0.07)		-0.07 (0.07)
Black Teacher		-0.14 (0.19)		-0.19 (0.19)		-0.04 (0.18)
Latinx/e Teacher		-0.12 (0.22)		-0.12 (0.22)		0.10 (0.20)
AAPI Teacher		0.11 (0.45)		-0.07 (0.44)		-0.10 (0.41)
Native Am. Teacher		-1.13 (0.76)		-2.25** (0.74)		-2.35** (0.70)
Multiracial Teacher		-0.48 (0.45)		-0.88* (0.44)		-1.35** (0.41)
Metro Area SES		0.14 (0.18)		0.03 (0.17)		0.04 (0.16)
Math Achievement		-0.41 (0.45)		-0.50 (0.43)		0.17 (0.41)
Avg. W-B/L Segregation		0.19 (0.53)		0.55 (0.52)		1.14* (0.49)
	R ² = 0.01	R ² = 0.05	R ² = 0.003	R ² = 0.09	R ² = 0.00	R ² = 0.17

*** $p < .001$, ** $p < 0.01$, * $p < 0.05$, † $p < 0.10$

Note: Coefficients are shown with standard errors in parentheses. $N=245$ for all unconditional models and $N=217$ for all conditional models due to observations missing covariate data from SEDA crosswalk. AAPI stands for Asian American and Pacific Islander. *Metro Area SES* is a composite score describing the socioeconomic status of school context metro area. *Math Achievement* is the cohort scale mean math achievement in the school context metro area. *Avg. W-B/L Segregation* represents the average of the school context metro area White-Black and White-Hispanic segregation indices relative to the local population, from 0 to 1, with 1 being total segregation. For instruction and attribution measures, we interpret coefficients as standardized differences in factor scores.

RQ3: How does teacher support for remediation, routinization, and extension activities in their next instructional move relate to their internal and external causal attributions for the student's struggle? Do these relationships differ across the conditions?

Looking at the relationship between teachers' positive attributions for the student's struggle and their instructional preferences, we find evidence that teachers' positive internal causal attributions—but not their external—predict their preferences for remediating, routinizing and extending the task. In other words, teacher endorsement of statements attributing the student struggle to positive internal factors, such as innate ability and effort—but not to external factors, such as instructional support and resources at home—positively predicted their preferences to reduce *and* extend the cognitive demand of the task. Teachers who more often attributed student performance to positive factors internal to the student were more likely to endorse activities that remediated ($\beta = 0.24, p = 0.091$), routinized ($\beta = 0.31, p = 0.019$), and extended the task ($\beta = 0.28, p = 0.031$). See Tables 2.7 and 2.8 for a summary of teachers' instructional preferences as they relate to their internal and external causal attributions, respectively.

Table 2.7. Teachers' Instructional Preferences on Positive Internal Causal Attributions

	Remediation		Routinization		Extension	
Internal Causal Attrib.	0.10 (0.11)	0.24 [†] (0.14)	0.24* (0.11)	0.31* (0.13)	0.18 (0.11)	0.28* (0.13)
Black Student	0.15 (0.15)	0.10 (0.18)	0.06 (0.15)	0.13 (0.17)	-0.02 (0.16)	0.11 (0.17)
Latinx/e Student	0.17 (0.16)	0.24 (0.18)	-0.11 (0.15)	0.04 (0.17)	-0.02 (0.16)	0.16 (0.17)
Internal X Black	0.01 (0.15)	-0.04 (0.18)	0.02 (0.15)	0.03 (0.17)	-0.19 (0.15)	-0.20 (0.17)
Internal X Latinx/e	0.15 (0.17)	0.08 (0.20)	0.04 (0.16)	0.00 (0.19)	-0.09 (0.17)	-0.19 (0.18)

Table 2.7. Continued

	Remediation		Routinization		Extension	
Passed Manip. Check		-0.18 (0.15)		-0.11 (0.14)		0.01 (0.14)
Grade Level Taught		0.03 (0.04)		0.04 (0.03)		0.08* (0.03)
Years Teaching		0.00 (0.01)		0.00 (0.01)		0.00 (0.01)
Woman		0.14 (0.15)		0.14 (0.14)		0.06 (0.14)
Non-Binary		0.75 (0.58)		0.34 (0.55)		0.55 (0.53)
MKT Score		-0.03 (0.07)		0.03 (0.07)		-0.07 (0.07)
Black Teacher		-0.18 (0.19)		-0.25 (0.18)		-0.05 (0.18)
Latinx/e Teacher		-0.19 (0.22)		-0.20 (0.21)		0.05 (0.21)
AAPI Teacher		-0.04 (0.45)		-0.30 (0.42)		-0.20 (0.41)
Native Am. Teacher		-1.23 (0.75)		-2.40** (0.71)		-2.38** (0.70)
Multiracial Teacher		-0.51 (0.45)		-0.95* (0.42)		-1.38** (0.41)
Metro Area SES		0.19 (0.17)		0.07 (0.16)		0.07 (0.16)
Math Achievement		-0.39 (0.44)		-0.45 (0.42)		0.16 (0.41)
Avg. W-B/L Segregation		0.12 (0.53)		0.48 (0.50)		1.01* (0.49)
	R ² = 0.03	R ² = 0.09	R ² = 0.07	R ² = 0.17	R ² = 0.01	R ² = 0.19

*** $p < .001$, ** $p < 0.01$, * $p < 0.05$, † $p < 0.10$

Table 2.7. Continued

Note: Coefficients are shown with standard errors in parentheses. $N=245$ for all unconditional models and $N=217$ for all conditional models due to observations missing covariate data from SEDA crosswalk. *Internal X Black* and *Internal X Latinx/e* represent interaction effect terms between teachers' internal causal attributions and racial-ethnic experimental condition, respectively. AAPI stands for Asian American and Pacific Islander. *Metro Area SES* is a composite score describing the socioeconomic status of school context metro area. *Math Achievement* is the cohort scale mean math achievement in the school context metro area. *Avg. W-B/L Segregation* represents the average of the school context metro area White-Black and White-Hispanic segregation indices relative to the local population, from 0 to 1, with 1 being total segregation. For instruction and attribution measures, we interpret coefficients as standardized differences in factor scores.

Table 2.8. Teachers' Instructional Preferences on Positive External Causal Attributions

	Remediation		Routinization		Extension	
External Causal Attrib.	0.05 (0.11)	0.06 (0.14)	0.18 (0.11)	0.16 (0.14)	0.03 (0.11)	0.04 (0.13)
Black Student	0.15 (0.16)	0.18 (0.18)	0.06 (0.15)	0.12 (0.17)	-0.01 (0.16)	0.08 (0.17)
Latinx/e Student	0.20 (0.16)	0.24 (0.19)	-0.09 (0.16)	0.03 (0.18)	-0.01 (0.16)	0.14 (0.17)
External X Black	0.00 (0.15)	0.02 (0.19)	0.03 (0.15)	0.07 (0.18)	0.11 (0.15)	0.14 (0.17)
External X Latinx/e	0.04 (0.16)	0.06 (0.20)	-0.03 (0.16)	0.01 (0.19)	0.10 (0.16)	0.05 (0.18)
Passed Manip. Check		-0.13 (0.15)		-0.04 (0.14)		0.04 (0.14)
Grade Level Taught		0.02 (0.04)		0.03 (0.03)		0.08* (0.03)
Years Teaching		0.00 (0.01)		0.00 (0.01)		0.00 (0.01)
Woman		0.08 (0.15)		0.06 (0.14)		0.01 (0.14)
Non-Binary		0.80 (0.59)		0.45 (0.57)		0.70 (0.54)
MKT Score		-0.01 (0.07)		0.05 (0.07)		-0.06 (0.07)

Table 2.8. Continued

	Remediation		Routinization		Extension	
Black Teacher		-0.16 (0.20)		-0.24 (0.19)		-0.08 (0.18)
Latinx/e Teacher		-0.14 (0.23)		-0.17 (0.22)		0.07 (0.21)
AAPI Teacher		0.08 (0.45)		-0.15 (0.43)		-0.18 (0.41)
Native Am. Teacher		-1.16 (0.77)		-2.29** (0.73)		-2.37** (0.70)
Multiracial Teacher		-0.50 (0.46)		-0.96* (0.44)		-1.40** (0.42)
Metro Area SES		0.15 (0.18)		0.03 (0.17)		0.04 (0.16)
Math Achievement		-0.43 (0.45)		-0.51 (0.43)		0.16 (0.41)
Avg. W-B/L Segregation		0.16 (0.54)		0.48 (0.51)		1.11* (0.49)
	R ² = 0.01	R ² = 0.05	R ² = 0.04	R ² = 0.13	R ² = 0.01	R ² = 0.18

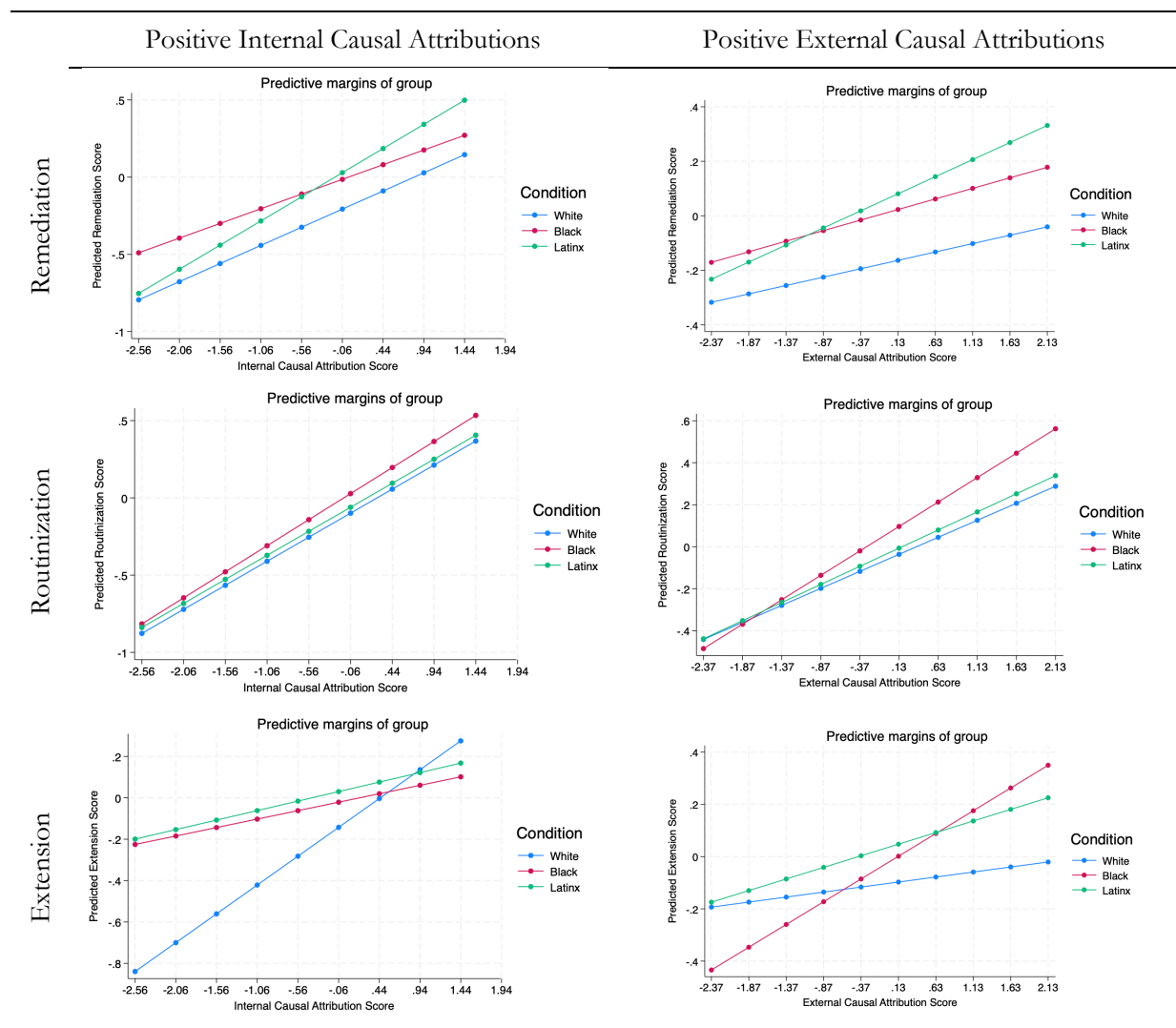
*** $p < .001$, ** $p < 0.01$, * $p < 0.05$, † $p < 0.10$

Note: Coefficients are shown with standard errors in parentheses. $N=245$ for all unconditional models and $N=217$ for all conditional models due to observations missing covariate data from SEDA crosswalk. *External X Black* and *External X Latinx/e* represent interaction effect terms between teachers' external causal attributions and racial-ethnic experimental condition, respectively. AAPI stands for Asian American and Pacific Islander. *Metro Area SES* is a composite score describing the socioeconomic status of school context metro area. *Math Achievement* is the cohort scale mean math achievement in the school context metro area. *Avg. W-B/L Segregation* represents the average of the school context metro area White-Black and White-Hispanic segregation indices relative to the local population, from 0 to 1, with 1 being total segregation. For instruction and attribution measures, we interpret coefficients as standardized differences in factor scores.

Examining the interaction effects between teachers' positive causal attributions and racial-ethnic condition, we do not any statistically significant differences. This tells us the slopes describing

the predicted relationship between teachers' causal internal and external attributions and their agreement with instructional approaches that remediate, routinize, and extend the math task did not differ whether teachers were confronted with a hypothetical struggling Black, White, or Latinx/e student. See Figure 2.2 for visualizations of the predicted relationship between teachers' positive causal internal and external attributions and their agreement with the three instructional approaches.

Figure 2.2. Relationship between teachers' internal causal attributions and instructional approaches by condition



Discussion

Relating teachers' internal and external causal attributions to student race-ethnicity, teacher characteristics and school context factors

I did not observe racial-ethnic biases in the ways teachers attributed the student's struggle in mathematics, either externally or internally. Relating teacher and school context factors to attributions, teachers who are women may be less likely to endorse positive internal causal attributions for the student struggle on the task as compared to men. Given that all students featured in the vignettes were girls, and prior work on attributional gender bias has shown girls' failures in math are more likely attributed to ability as compared to boys (e.g. Espinoza et al., 2014), it could be the case that there is an interaction effect between teacher and student gender on attributions. More work is needed to understand the role of student-teacher gender match on teachers' causal attributions in mathematics.

Moreover, we find trend-level evidence to suggest teachers of more advanced grade levels may be less likely to explain the student struggle with positive internal factors compared to teachers of earlier grade levels. This emergent finding extends inconsistent findings by Wang and Hall (2018) relating level of instruction and teachers' causal attributions. It could be the case that teachers of earlier grades may be more likely to attribute student challenges to the student given their relatively recent entry into the school system. Additionally, those teaching more advanced grade levels may also be more likely to support extending the math task. Level of instruction corresponds with teachers' content knowledge of the elementary mathematics curriculum, and those with greater knowledge may be more willing to move students towards advanced topics due to their content fluency.

Relating teacher support for task remediation, routinization, and extension to student race-ethnicity, teacher characteristics and school context factors

Moreover, I did not observe racial-ethnic biases in teacher preferences to remediate, routinize, or extend the math task. Here, the focus on mathematics content rather than the student may have served as a leveling effect, as prior research shows that an emphasis on objective information such as content rather than student identity may diminish the potential for biases to inform judgements and decision making (e.g. Quinn, 2020).

Looking at school context factors, the level of segregation in the metro area relative to the local population related to teacher agreement to extend the task. The greater the average White-Black and White-Hispanic school segregation relative to the local population, the more likely teachers agreed to extend the task. Somewhat paradoxically, this study found that teachers' instructional preferences did not differ by racial-ethnic condition but did differ descriptively according to the level of racial-ethnic segregation in the teaching context. Prior research on ambitious math teaching has shown that salient dimensions of student identity such as race may shape teachers' attributionally informed instructional practice, with a stronger relationship for classrooms composed of historically marginalized learners (Wilhelm et al., 2017). The present study builds on these findings by taking social context into account, which has been theoretically and empirically linked to racial-ethnic biases (Chin et al., 2020; Payne et al., 2017). It could be the case that in contexts in which racially minoritized students and White students do not attend school together, student race becomes less salient, playing less of a role in shaping teacher preferences to increase cognitive demand. More research is needed to understand how contextual racial segregation shapes teachers' attributions and instructional practice in math, with implications for teacher development in service of greater equity in mathematics education.

Relating teachers' causal attributions and their preferences to alter cognitive demand

In the present study, I found teacher endorsement of statements attributing the student struggle to positive internal factors, such as innate ability and effort—but not to external factors, such as instructional support and resources at home—positively related to their preferences to both reduce *and* extend the cognitive demand of the task. In other words, the more strongly teachers agreed that the student was engaged in class, mathematically talented, and motivated, the more likely they were to adapt the task. However, these same task adjustments were not predicted by teacher agreement with positive external attributions (that the student enjoyed math resources at home, received strong prior math instruction, and had prior experience with challenging math tasks). These finding shows that teacher preferences to take instructional action (whether it be to decrease or increase the level of cognitive demand) relates to where they locate the source of student difficulties in math. When teachers can explain student difficulties in terms of their effort, engagement, and natural ability— rather than things outside of the locus of the student, such as prior instruction and support at home— this predicts teacher preferences to remediate, routinize, and extend the task. These attribution-instructional preference relationships were comparable whether teachers were confronted with a hypothetical struggling Black, White, or Latinx/e student.

These findings extend those by Wilhelm and colleagues (2017) which showed teachers who explained student difficulties in mathematics as a function of their instructional experiences orchestrated greater opportunities for them to share their mathematical reasoning in class, compared to classrooms with teachers who explained difficulties as stemming from perceived deficits inherent to them or their communities. In other words, when teachers believed student difficulties stemmed from perceived deficits, teachers may have felt their instructional choices were futile. On the other hand, teachers that felt instruction caused these difficulties saw opportunity to learn as an important

lever (Wilhelm et al., 2017). Echoing some of Wilhelm et al.'s (2017) findings, in the present study, the view that student difficulties stem from external factors did not relate to instructional preferences. This finding could similarly be due to a view that external factors are “out of their hands,” whereas a locus within the student positively informs instruction because it is more actionable.

Limitations

There are several limitations to the present study. First, all teachers responded to the work of a hypothetical rather than actual student. Given that in actuality, teachers make attributions about students whom they know and with whom they have relationships, the phenomena of teacher' attributions and instructional preferences are best understood in context. Additionally, this study measured teachers' reported preferences rather than their actions in the classroom with respect to cognitive demand. Due to the experimental nature of this study, there were tradeoffs between the internal and ecological validity. Finally, teachers' attributions were measured using Likert-type items from the literature on internal and external causal attributions. A more comprehensive measure would have disentangled teachers' attributions along dimensions of stability and control as well as locus. For example, the Attributions of Mathematical Excellence Scale (AMES) developed by Jacobson et al. (2022) measures genetic, social, personal, and educational attributions. However, given the limited survey length required to achieve a viable sample, this presented challenges and trade-offs. Finally, all attributions items were worded positively as opposed to neutrally. While scholars have used items worded neutrally negatively, and positively to measure teachers' attributions (Copur-Gencturk et al., 2024; Jacobson et al., 2022), future work in this area may adhere to a mixture of items for measuring teachers' attributions.

Conclusion

The present study found teachers' internal causal attributions about student struggles in math positively related to their preferences to take instructional action, whether it was to remediate, routinize, or extend the task. Further, teachers were found to be equitable in the ways they explained and responded to student stuck points, suggesting relatively equitable practice in this area of ambitious math teaching. Because teacher preferences to alter the cognitive demand of the task related to their internal rather than external causal attributions for the student struggle, teacher practices that diagnose student-facing misunderstandings, such as via formative assessment, are an important lever for informing ensuing practice.

References

- Bobis, J. (2024). Core Practices That Support Learning to Teach Mathematics Through a Challenging Problem-Solving Approach. In P. L. Grossman & U. Fraefel (Eds.), *Core practices in teacher education: A global perspective*. (pp. 117-138). Harvard Education Press.
- Borasi, R. (1994). Capitalizing on Errors as “Springboards for Inquiry”: A Teaching Experiment. *Journal for Research in Mathematics Education*, 25(2), 166–208. <https://doi.org/10.2307/749507>
- Boston, M. (2012). Assessing Instructional Quality in Mathematics. *The Elementary School Journal*, 113(1), 76–104. <https://doi.org/10.1086/666387>
- Chin, M. J., Quinn, D. M., Dhaliwal, T. K., & Lovison, V. S. (2020). Bias in the Air: A Nationwide Exploration of Teachers’ Implicit Racial Attitudes, Aggregate Bias, and Student Outcomes. *Educational Researcher*, 49(8), 566–578. <https://doi.org/10.3102/0013189X20937240>
- Copur-Gencturk, Y., Thacker, I., & Cimpian, J. R. (2024). An exploratory experiment investigating teachers’ attributional race and gender bias and the moderating effects of personal experience of racial discrimination. *Contemporary Educational Psychology*, 79, 102317. <https://doi.org/10.1016/j.cedpsych.2024.102317>
- Copur-Gencturk, Y., Thacker, I., & Cimpian, J. R. (2022). Teacher bias in the virtual classroom. *Computers & Education*, 191, 104627. <https://doi.org/10.1016/j.compedu.2022.104627>
- Dobbs, J., & Arnold, D. H. (2009). The relationship between preschool teachers’ reports of children’s behavior and their behavior towards those children. *School Psychology Quarterly*, 24(2), 95–105.
- Dovidio, J. F. (2001). On the Nature of Contemporary Prejudice: The Third Wave. *Journal of Social Issues*, 57(4), 829–849. <https://doi.org/10.1111/0022-4537.00244>
- Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Harcourt Brace Jovanovich College Publishers.
- Espinoza, P., Arêas da Luz Fontes, A. B., & Arms-Chavez, C. J. (2014). Attributional gender bias: Teachers’ ability and effort explanations for students’ math performance. *Social Psychology of Education*, 17(1), 105-126.
- Estrella, S., Zakaryan, D., Olfos, R., & Espinoza, G. (2020). How teachers learn to maintain the cognitive demand of tasks through Lesson Study. *Journal of Mathematics Teacher Education*, 23(3), 293–310. <https://doi.org/10.1007/s10857-018-09423-y>
- Fahle, E. M., Reardon, S. F., Shear, B. R., Ho, A. D., Saliba, J., Min, J., Shim, J. & Kalogrides, D. (2025). Stanford Education Data Archive Technical Documentation: SEDA 2024. Retrieved from <https://purl.stanford.edu/pt329xg7054>.
- Froehlich, L., Martiny, S. E., Deaux, K., & Mok, S. Y. (2016). “It’s their responsibility, not ours”: Stereotypes about competence and causal attributions for immigrants’ academic underperformance. *Social Psychology*, 47(2), 74–86. <https://doi.org/10.1027/1864-9335/a000260>

- Gaddis, S. M. (2017a). How Black Are Lakisha and Jamal? Racial Perceptions from Names Used in Correspondence Audit Studies. *Sociological Science*, 4(19), 469–489. <https://doi.org/10.15195/v4.a19>
- Gaddis, S. M. (2017b). Racial/Ethnic Perceptions from Hispanic Names: Selecting Names to Test for Discrimination. *Socius*, 3. <https://doi.org/10.1177/2378023117737193>
- Georgiou, S. N. (2008). Beliefs of experienced and novice teachers about achievement. *Educational Psychology*, 28(2), 119–131. <https://doi.org/10.1080/01443410701468716>
- Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin*, 117(1), 21–38. <https://doi.org/10.1037/0033-2909.117.1.21>
- Graham, S. (1991). A review of attributions theory in achievement contexts. *Educational Psychology Review*, 3, 5–39.
- Gottfried, M., Fletcher, T., & Comstock, M. (2023). Teaching mathematics in kindergarten: How does instruction differ by classroom ability? *Teachers College Record*, 125(10), 102–130. <https://doi.org/10.1177/01614681231223343>
- Henningsen, M., & Stein, M. K. (1997). Mathematical tasks and student cognition: Classroom-based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for Research in Mathematics Education*, 28(5), 524–549.
- Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In J. Frank & K. Lester (Eds.), *Second handbook of research on mathematics teaching and learning* (pp. 371–404). Charlotte: Information Age.
- Hill, H. C., Lovison, V., & Kelley-Kemple, T. (2019). Mathematics Teacher and Curriculum Quality, 2005 and 2016. *AERA Open*, 5(4).
- Hill, H. C., Ball, D. L., Blunk, M., Goffney, I. M., & Rowan, B. (2007). Validating the ecological assumption: The relationship of measure scores to classroom teaching and student learning. *Measurement*, 5(2-3), 107-118
- Hilton, J. L., & von Hippel, W. (1996). Stereotypes. *Annual Review of Psychology*, 47, 237–271. <https://doi.org/10.1146/annurev.psych.47.1.237>
- Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A Meta-Analysis on the Correlation Between the Implicit Association Test and Explicit Self-Report Measures. *Personality & Social Psychology Bulletin*, 31(10), 1369–1385. <https://doi.org/10.1177/0146167205275613>
- Jackson, K., Gibbons, L., & Sharpe, C. J. (2017). Teachers' Views of Students' Mathematical Capabilities: Challenges and Possibilities for Ambitious Reform. *Teachers College Record*, 119(7), 1–43. <https://doi.org/10.1177/016146811711900708>
- Jacobson, E., Cross Francis, D., Willey, C., & Wilkins-Yel, K. (2022). Race, Gender, and Teacher Equity Beliefs: Construct Validation of the Attributions of Mathematical Excellence Scale. *AERA Open*, 8, 23328584221130999. <https://doi.org/10.1177/23328584221130999>

- Jussim, L., & Harber, K. D. (2005). Teacher expectations and self-fulfilling prophecies: knowns and unknowns, resolved and unresolved controversies. *Personality and Social Psychology Review*, 9, 131-155. DOI:10.1207/s15327957pspr0902_3
- Kline, R. B. (2016). *Principles and Practice of Structural Equation Modeling* (4th ed.). Guilford Press.
- Lampert, M., Franke, M. L., Kazemi, E., Ghouseini, H., Turrou, A. C., Beasley, H., Cunard, A., & Crowe, K. (2013). Keeping it complex: Using rehearsals to support novice teacher learning of ambitious teaching. *Journal of Teacher Education*, 64(3), 226–243. <https://doi.org/10.1177/0022487112473837>
- Martin, D. B. (2009). Researching Race in Mathematics Education. *Teachers College Record*, 111(2), 295-338. <https://doi-org.ezp-prod1.hul.harvard.edu/10.1177/016146810911100208>
- Middleton, J. A., & Spanias, P. (1999). Motivation for achievement in mathematics: Findings, generalizations, and criticisms of the research. *Journal for Research in Mathematics Education*, 30, 65–88.
- National Center for Education Statistics (NCES). (2023). Characteristics of Public School Teachers. *Condition of Education*. U.S. Department of Education, Institute of Education Sciences. Retrieved from <https://nces.ed.gov/programs/coe/indicator/clr>.
- National Research Council. (1991). *Moving beyond myths: Revitalizing undergraduate mathematics*. Washington, DC: National Academy Press.
- Owens, A., Reardon, S.F., Kalogrides, D., Jang, H., and Tom, T. (2022). Measuring Segregation with the Normalized Exposure Index. Research Brief. *The Segregation Index*.
- Payne, B. K., Vuletich, H. A., & Lundberg, K. B. (2017). The Bias of Crowds: How Implicit Bias Bridges Personal and Systemic Prejudice. *Psychological Inquiry*, 28(4), 233–248. <https://doi.org/10.1080/1047840X.2017.1335568>
- Peterson, E. R., Rubie-Davies, C., Osborne, D., & Sibley, C. (2016). Teachers’ explicit expectations and implicit prejudiced attitudes to educational achievement: Relations with student achievement and the ethnic achievement gap. *Learning and Instruction*, 42, 123–140. <https://doi.org/10.1016/j.learninstruc.2016.01.010>
- Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, 90(5), 751–783. <https://doi.org/10.1037/0022-3514.90.5.751>
- Pirrone C. (2012). The influence of teachers’ preexisting notions about students on scholastic achievement. *AASA Journal of Scholarship and Practice*, 9, 18–28.
- Pollard, C. (2022). Teaching, Learning, and Race: Toward the Identification of Mechanisms Underlying Own- and Other-Race Teacher Effects. Doctoral dissertation, Harvard University Graduate School of Arts and Sciences
- Quinn, D. M. (2020). Experimental Evidence on Teachers’ Racial Bias in Student Evaluation: The Role of Grading Scales. *Educational Evaluation and Policy Analysis*, 42(3), 375–392. <https://doi.org/10.3102/0162373720932188>

- Reyna, C. (2000). Lazy, Dumb, or Industrious: When Stereotypes Convey Attribution Information in the Classroom. *Educational Psychology Review*, 12(1), 85–110. <https://doi.org/10.1023/A:1009037101170>
- Reyna, C., & Weiner, B. (2001). Justice and utility in the classroom: An attributional analysis of the goals of teachers' punishment and intervention strategies. *Journal of Educational Psychology*, 93(2), 309–319. <https://doi.org/10.1037/0022-0663.93.2.309>
- Reyna, C. (2008). Ian is intelligent but Leshawn is lazy: Antecedents and consequences of attributional stereotypes in the classroom. *European Journal of Psychology of Education*, 23(4), 439–458. <https://doi.org/10.1007/BF03172752>
- Satorra, A. (2000). Scaled and adjusted restricted tests in multi-sample analysis of moment structures. In Heijmans, R.D.H., Pollock, D.S.G. & Satorra, A. (eds.), *Innovations in multivariate statistical analysis. A Festschrift for Heinz Neudecker* (pp.233-247). London: Kluwer Academic Publishers.
- Schneider, D. J. (2004). *The psychology of stereotyping*. Guilford Press.
- Schoenfeld, Alan & Fink, Heather & Zuniga Ruiz, Sandra & Huang, Siqi & Wei, Xinyu & Chirinda, Brantina. (2023). *Helping Students Become Powerful Mathematics Thinkers: Case Studies of Teaching for Robust Understanding*. 10.4324/9781003375197.
- Smith, M. S., & Stein, M. K. (1998). Reflections on Practice: Selecting and Creating Mathematical Tasks: From Research to Practice. *Mathematics Teaching in the Middle School*, 3(5), 344–350. <https://doi.org/10.5951/MTMS.3.5.0344>
- Stein, M. K., Smith, M. S., Henningsen, M. A., & Silver, E. A. (2009). *Implementing standards-based math instruction: A casebook for professional development*. Teachers College Press.
- Stein, M. K. & Smith, M. S. (1998). “Mathematical tasks as a framework for reflection: From research to practice.” *Mathematics teaching in the middle school* 3.4: 268-275.
- Sullivan, P., Askew, M., Cheeseman, J., Clarke, D. M., Mornane, A., Roche, A., & Walker, N. (2014). Supporting teachers in structuring mathematics lessons involving challenging tasks. *Journal of Mathematics Teacher Education*, 1-18.
- Sullivan, P., & Davidson, A. (2014). The role of challenging mathematical tasks in creating opportunities for student reasoning. In J. Anderson, M. Cavanagh & A. Prescott (Eds.), *Proceedings of the 37th annual conference of the Mathematics Education Research Group of Australasia* (pp. 605-612). Sydney, Australia: MERGA.
- Sullivan, P., Mousley, J., & Zevenbergen, R. (2006). Teacher Actions to Maximize Mathematics Learning Opportunities in Heterogeneous Classrooms. *International Journal of Science and Mathematics Education*, 4(1), 117–143. <https://doi.org/10.1007/s10763-005-9002-y>
- Tajfel, H. (1970). Experiments in Intergroup Discrimination. *Scientific American*, 223(5), 96–103.
- Thames, M. H., & Ball, D. L. (2010). What math knowledge does teaching require? *Teaching Children Mathematics*, 17(4), 220-229.

- VanLehn, K., Siler, S., Murray, C., Yamauchi, T., & Baggett, W. B. (2003). Why Do Only Some Events Cause Learning During Human Tutoring? *Cognition and Instruction*, 21(3), 209–249. https://doi.org/10.1207/S1532690XCI2103_01
- Wang, H. & Hall, N. C. (2018). A Systematic Review of Teachers' Causal Attributions: Prevalence, Correlates, and Consequences. *Frontiers in Psychology*, 9. <https://doi.org/10.3389/fpsyg.2018.02305>
- Warshauer, H. K. (2015). Productive struggle in middle school mathematics classrooms. *Journal of Mathematics Teacher Education*, 18(4), 375–400. <https://doi.org/10.1007/s10857-014-9286-3>
- Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, 92(4), 548–573. <https://doi.org/10.1037/0033-295X.92.4.548>
- Weiner, B. (2000). Intrapersonal and Interpersonal Theories of Motivation from an Attributional Perspective. *Educational Psychology Review* 12, 1–14 <https://doi.org/10.1023/A:1009017532121>
- Weiner, B. (2010). The Development of an Attribution-Based Theory of Motivation: A History of Ideas. *Educational Psychologist*, 45(1), 28–36. <https://doi.org/10.1080/00461520903433596>
- Wilhelm, A. G., Munter, C., & Jackson, K. (2017). Examining Relations between Teachers' Explanations of Sources of Students' Difficulty in Mathematics and Students' Opportunities to Learn. *The Elementary School Journal*, 117(3), 345–370. <https://doi.org/10.1086/690113>
- Wilson, P. H., Edgington, C., Sztajn, P., & DeCuir-Gunby, J. (2014). Teachers, Attributions, and Students' Mathematical Work. In J.-J. Lo, K. R. Leatham, & L. R. Van Zoest (Eds.), *Research Trends in Mathematics Teacher Education* (pp. 115–132). Springer International Publishing. https://doi.org/10.1007/978-3-319-02562-9_7

Appendix A2

Table A2.1. Content and features of each vignette

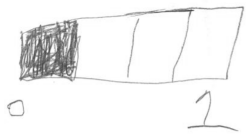
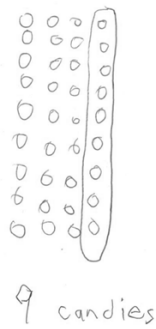
Vignette	Prompt	Mathematics Task and Student Work
1	Your class is working on the task below about identifying fractional parts on a number line. You notice a student has drawn a fraction diagram with four equal parts and $\frac{1}{4}$ shaded in out of 1 whole, but they have not labeled the number line, as shown on their paper below. What would you say to this student to support them in locating 1 on the number line?	[Condition A] Name <u>Tanisha Washington</u> Locate 1 on the number line. Be as exact as possible  
2	Your class is working on the task below about equivalent fractions using a brownie pan cut into 16 pieces. You notice a student correctly represented 16 brownies, but they confused $\frac{7}{8}$ of 16 for $\frac{7}{16}$, as shown on their paper below. What would you say to this student to support them in finding the number of brownies left?	[Condition B] Name <u>Katie W.</u> Lisa baked a large pan of brownies and cut them into 16 equal pieces. She sold $\frac{7}{8}$ of the brownies at the school bake sale. How many brownies does Lisa have left? 
3	Your class is working on the task below about dividing 36 candies evenly among 9 children. You notice a student has drawn a 4 x 9 array totaling 36 candies but incorrectly stated that each child would receive 9 candies, as shown on their paper below. What would you say to this student to support them in finding out how many pieces of candy each kid will get?	[Condition C] Name <u>Daniela</u> There are 36 pieces of candy that need to be evenly divided among nine kids. How many pieces of candy will each kid get? 

Table A2.1. Continued

4 Your class is working on the place value task below. You notice a student decided to take the sum of both numbers but incorrectly wrote 4 tens as the number 4 instead of 40, as shown on their paper below. What would you say to this student to support them in using place value to find the mystery number?	[Condition A] Name <u>Tanisha Washington</u> ? "I have 23 ones and 4 tens. What number am I?" ? $\begin{array}{r} 23 \\ + 4 \\ \hline 27 \end{array}$										
5 Your class is working on the task below about choosing an appropriate scale for a bar graph that represents a data set. You notice a student has correctly interpreted the data in the table to create a bar graph, but they have chosen a scale of one, which limits the range of values they are able to represent on the graph, as shown on their paper below. What would you say to this student to support them in choosing an appropriate scale for their bar graph?	[Condition B] Name <u>Katie</u> All the third-grade students at school were asked, "What is your favorite season?" Their responses are shown in this table. Use the data from this table to create a scaled bar graph. <table border="1"> <thead> <tr> <th>favorite time of the year</th><th>number of students</th></tr> </thead> <tbody> <tr> <td>winter</td><td>24</td></tr> <tr> <td>spring</td><td>13</td></tr> <tr> <td>summer</td><td>40</td></tr> <tr> <td>fall</td><td>22</td></tr> </tbody> </table>	favorite time of the year	number of students	winter	24	spring	13	summer	40	fall	22
favorite time of the year	number of students										
winter	24										
spring	13										
summer	40										
fall	22										
6 Your class is working on the task below about representing and solving addition and subtraction problems. You notice a student correctly understood they need to subtract the length of 27, but they have not yet found the missing value, as shown on their paper below. What would you say to this student to support them in finding the missing value in the diagram?	[Condition C] Name <u>Daniela Velasquez</u> Find out the missing value in the diagram below (assume both long bars are equal in total length). 48 - 27										

Table A2.1. Continued

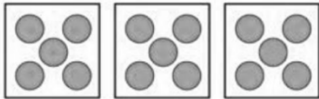
7 Your class is working on the multiplication task below. You notice a student correctly found the total number of marbles, but they did not use a multiplication sentence, as shown on their paper below. What would you say to this student to support them in using a multiplication sentence to find out how many marbles they will buy in total?	[Condition A] Name <u>Tanisha W.</u> The toy store is selling boxes of marbles. Each box comes with five marbles. A shopper wants to buy three boxes. Use a <u>multiplication sentence</u> to find out how many marbles they will buy in total.  $\begin{array}{r} 5 \\ + 5 \\ \hline 10 \\ + 5 \\ \hline 15 \end{array}$
8 Your class is working on the task below about subtracting dollars and cents. You notice a student correctly chose subtraction but did not account for the correct place value, as shown on their paper below. What would you say to this student to support them in subtracting \$9.50 from \$135?	[Condition B] Name <u>Katie W.</u> Jodie saved up \$135 dollars from her summer job. She spent \$9.50 on a movie ticket. How much money does she have left? $\begin{array}{r} 0 \\ 135 \\ - 9.50 \\ \hline \$ 9.85 \end{array}$

Table A2.2. Items measuring teacher preference for task remediation

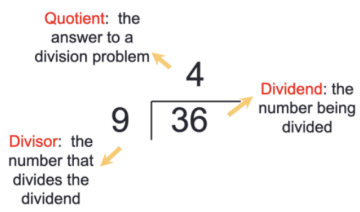
Vignette	Item
1	Instead of having the student continue to engage in this task about identifying fractional parts on a number line, I would prioritize having them practice something more straightforward. There are several ways that this could be productively done, which might include identifying fractional parts on a fraction diagram (e.g. $\frac{1}{4}$ of a shape is shaded), or some other activity in this vein.
2	Instead of having the student continue to engage in this task about equivalent fractions, I would prioritize having them practice something more straightforward. There are several ways that this could be productively done, which might include having them simplify the fraction $14/16$ by dividing 14 and 16 by 2 to get $7/8$, or some other activity in this vein.
3	Instead of having the student continue to engage in this division task, I would prioritize having them practice something more straightforward. There are several ways that this could be productively done, which might include having the student set up and label the parts of a division problem (as shown below), or some other activity in this vein. 
4	Instead of having the student continue to engage in this place value task, I would prioritize having them practice something more straightforward. There are several ways that this could be productively done, which might include having the student identify the place value of numbers (e.g. <i>In the number 34, we have <u>3</u> in the tens place and <u>4</u> in the ones place</i>), or some other activity in this vein.
5	Instead of having the student continue to engage in this task about creating a scaled bar graph, I would prioritize having them practice something more straightforward. There are several ways that this could be productively done, which might include having the student figure out if they should count by 2's, 5's or 10's for a given dataset, or some other activity in this vein.
6	Instead of having the student continue to engage in this task about representing and solving addition and subtraction problems, I would prioritize having them practice something more straightforward. There are several ways that this could be productively done, which might include having the student solve addition and subtraction problems that have already been set up, or some other activity in this vein.

Table A2.2. Continued

Instead of having the student continue to engage in this multiplication task, I would prioritize having them practice something more straightforward. There are several ways that this could be productively done, which might include having the student compare repeated addition and multiplication (as shown below), or some other activity in this vein.

	$2 + 2 + 2$ $= \underline{\hspace{2cm}}$	$3 \times 2 = \underline{\hspace{2cm}}$
	$3 + 3 + 3$ $= \underline{\hspace{2cm}}$	$3 \times 3 = \underline{\hspace{2cm}}$
	$4 + 4$ $= \underline{\hspace{2cm}}$	$2 \times 4 = \underline{\hspace{2cm}}$

Instead of having the student continue to engage in this subtraction task, I would prioritize having them practice something more straightforward. There are several ways that this could be productively done, which might include having the student set up problems by writing in and aligning the decimal points (e.g. \$5 $\rightarrow$ \$5.00), or some other activity in this vein.

Table A2.3. Items measuring teacher preference for task routinization

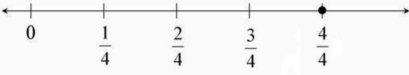
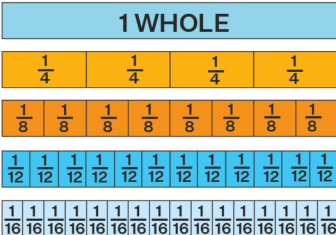
Vignette	Item
1	While the student is working on this task, I would provide them with more direct support. There are several ways that this could be productively done, which might include labeling the entire number line for the student so they know where $\frac{1}{4}$, $\frac{2}{4}$, $\frac{3}{4}$, and $\frac{4}{4}$ are (as shown below), or some other activity in this vein. 
2	While the student is working on this task, I would provide them with more direct support. There are several ways that this could be productively done, which might include providing a visual anchor chart to help the student convert between unlike fractions (as shown below), or some other activity in this vein. 

Table A2.3. Continued

3 While the student is working on this task, I would provide them with more direct support. There are several ways that this could be productively done, which might include setting up the division problem 36 divided by 9 (as shown below) or some other activity in this vein.

$$9 \overline{) 36}$$

4 While the student is working on this task, I would provide them with more direct support. There are several ways that this could be productively done, which might include having the student place the numbers in a place value chart (as shown below), or some other activity in this vein.

	Tens	Ones	
+			
=			
	Tens	Ones	Number

5 While the student is working on this task, I would provide them with more direct support. There are several ways that this could be productively done, which might include providing a pre-labeled graph that counts by 5's on the y-axis (as shown below), or some other activity in this vein.

6 While the student is working on this task, I would provide them with more direct support. There are several ways that this could be productively done, which might include providing the student with the total length of the bottom bar, i.e. showing 83 rather than 48+35, or some other activity in this vein.

7 While the student is working on this task, I would provide them with more direct support. There are several ways that this could be productively done, which might include providing blanks for the student to fill in (as shown below), or some other activity in this vein.

of boxes X # of marbles per box = Total # of marbles

_____ X _____ = _____

8 While the student is working on this task, I would provide them with more direct support. There are several ways that this could be productively done, which might include using play money and coins to count out the difference, or some other activity in this vein.

Table A2.4. Items measuring teacher preference for task extension

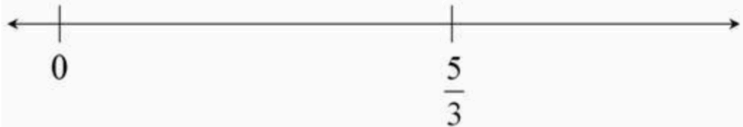
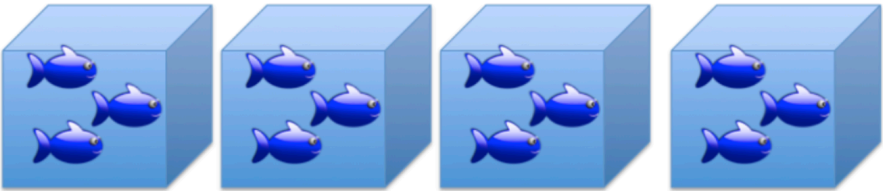
Vignette	Item
1	After supporting this student on this stuck-point, I would have them practice locating 1 on the number line again with a slightly more challenging task. This next task could take many forms, and an example of what one could look like is shown below: Locate 1 on the number line. Label the point and be as exact as possible. 
2	After supporting this student on this stuck-point, I would have them practice equivalent fractions again with a slightly more challenging task. This next task could take many forms, and an example of what one could look like is shown below: Kenny made a cake for the party and cut it into 20 slices. By the end of the party, $\frac{3}{4}$ of the cake was gone. How many slices were left?
3	After supporting this student on this stuck-point, I would have them practice division again with a slightly more challenging task. This next task could take many forms, and an example of what one could look like is shown below: Suppose there are 4 tanks and 3 fish per tank. The total number of fish in this situation can be expressed as $4 \times 3 = 12$  1) Describe what is meant in this situation by $12 \div 3 = 4$ 2) Describe what is meant in this situation by $12 \div 4 = 3$
4	After supporting this student on this stuck-point, I would have them practice place value again with a slightly more challenging task. This next task could take many forms, and an example of what one could look like is shown below: I have 7 tens and 43 ones. What number am I?

Table A2.4. Continued

After supporting this student on this stuck-point, I would have them practice creating a scaled bar graph again with a slightly more challenging task. This next task could take many forms, and an example of what one could look like is shown below:

5

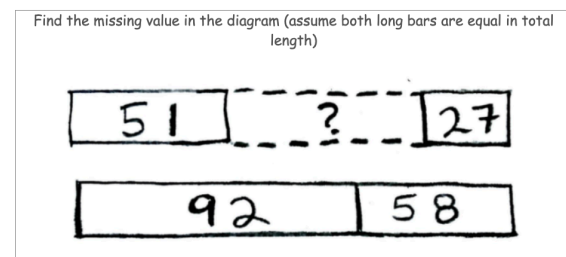
Create a scaled bar graph for the dataset below:

Year	Rain (inches)
2021	103
2022	38
2023	87
2024	14



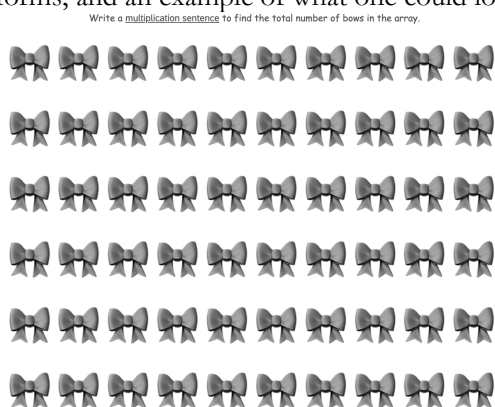
After supporting this student on this stuck-point, I would have them practice representing and solving addition and subtraction problems again with a slightly more challenging task. This next task could take many forms, and an example of what one could look like is shown below:

6



After supporting this student on this stuck-point, I would have them practice multiplication again with a slightly more challenging task. This next task could take many forms, and an example of what one could look like is shown below:

7



After supporting this student on this stuck-point, I would have them practice subtraction with money again with a slightly more challenging task. This next task could take many forms, and an example of what one could look like is shown below:

8

The school bake sale raised \$347.75 after spending \$52 on ingredients. How much money did the school profit from the bake sale?

Chapter 3

Not Having All the Answers: Connections, *Doing Math*, and Teacher Preservation of Task Cognitive Demand in Response to a Student Fractions Misconception

Abstract

Students in the U.S. show enduring misconceptions about fractions, curtailing their progression in mathematics. While the field has some understanding about how teachers respond to and resolve student misconceptions on proportional reasoning tasks (e.g. Warshauer, 2015), it lacks empirical work mapping out teacher treatment of math content and degree of student accountability when implementing a high-level fractions task in the face of student challenges. Understanding teacher treatment of fractions misconceptions is important in light of historic reforms pushing for cognitively demanding mathematics instruction and meaning-centered mathematics curricula that centers student understandings, particularly in the fraught topic of fractions. In this study, I solicit teacher responses to a student misconception about fractions via an instructional vignette, finding that most teachers forged at least one connection between fractions concepts, procedures, and their underlying meaning, namely through connecting the problem context and visual fraction diagram. Still, a large portion of teachers focused on the steps of the standard algorithm without mathematical connections, and over half took a teacher-directed problem-solving approach. Notably, scaffolding and probing often co-occurred with privileging the student strategy, suggesting a potentially complimentary role between these approaches that enlist the student in problem solving. The findings have implications for research on cognitive demand in the implementation of math tasks and improving teacher knowledge, practice, and student learning in the foundational domain of fractions.

Not Having All the Answers: Connections, *Doing Math*, and Teacher Preservation of Task Cognitive Demand in Response to a Student Fractions Misconception

Both the Common Core State Standards for Mathematics (CCSSM) and National Council of Teachers of Mathematics (NCTM) outline the importance of mastering concepts like fractional equivalence starting in grade 3 (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010; NCTM, 2007). Despite these goals for fractions learning, a concerning proportion of elementary students in the U.S. make little progress, demonstrating enduring misconceptions and shallow understandings of fractions (Hansen et al., 2017; Mack, 2001). Common challenges and misconceptions about fractions include the (mis)transfer of whole number concepts, seeing numerators and denominators as separate values rather than in relation, and viewing fractions and fractions operations as the arbitrary manipulation of symbols absent of meaning (Siegler et al., 2010). Teachers, too, have been shown to hold fragmented knowledge of fractions (Copur-Gencturk, 2021; Ma, 2006; Newton, 2008; Olanoff et al., 2014), and elementary fractions instruction in the U.S. often narrowly focuses on representing fractions with part-whole diagrams and practice with fractions operations absent of connections between concepts, procedures, and their underlying meanings (Gearhart et al., 1999; Siegler et al., 2010). In short, “no area of elementary school mathematics is as mathematically rich, cognitively complicated, and difficult to teach as fractions, ratios, and proportionality” (Smith, 2002, p. 3)

To support students in building enduring understandings of fractions and indeed mathematics more generally, scholars have argued for the use of cognitively challenging mathematical tasks in the classroom (Boston, 2012; Siegler et al., 2010; Stein et al., 2009; Stein & Smith, 1998). A math task is a piece of class work with an intended product, and a specified process for creating that product, using a particular set of resources under certain classroom conditions

(Doyle, 1988). The construction of a math task is intended to focus student thinking about a mathematical idea, determining (along with its enactment) how students come to think about a mathematical domain and its meaning (Doyle, 1988). Math tasks can be differentiated based on their cognitive level or the nature of the mathematical thinking processes they invoke (Doyle, 1988; Stein et al., 1996). In contrast to tasks with low-level cognitive demands, which involve reproducing memorized rules or procedures without connections, tasks with high-level cognitive demands are those that involve procedures *with connections* and students *doing* mathematics (Stein et al., 1996). Thus, cognitive demand describes the nature of cognitive processes students engage in, which “can range from memorization of procedures and algorithms (with or without attention to concepts, understanding, or meaning) to complex thinking and reasoning strategies that would be typical of ‘doing mathematics’ (e.g., conjecturing, justifying, or interpreting)” (Henningesen & Stein, 1997, p. 529; Schoenfeld et al., 2023). Following this line of scholarship, I define the cognitive demand of task-based mathematics instruction in terms of the connectedness of the mathematics involved, i.e. connections between procedures, concepts, and their underlying meanings, and the extent to which students are accountable for high-level mathematical processes and products (Boston, 2012; Henningesen & Stein, 1997; Hiebert & Grouws, 2007; Schoenfeld et al., 2023; Stein & Lane, 1996).

Despite the promise of tasks as intended, the implementation stage in the classroom often leads to the devolution of tasks with high-level cognitive demands, for example through routinization or remediation (Estrella et al., 2020; Henningesen & Stein, 1997). Moreover, student misconceptions about fractions during task implementation—which necessitate instructional responses that enable students to forge connections between concepts, procedures, and their underlying meaning, and to practice *doing* and *making sense of* mathematics in order to address them—may create impasses during problem solving that trigger a wide spectrum of instructional responses from teachers (Henningesen & Stein, 1997; Schoenfeld et al., 2023; Siegler et al., 2010; Stein & Smith,

1998; Warshauer, 2015). In these instances, the teacher may respond by simplifying or preserving the mathematical connections inherent to the task. They may also take over the problem solving or keep the student enlisted as an agentic doer of mathematics. Few studies have examined how teachers respond to inevitable, expected, and necessary student stuck points or impasses during the implementation of fractions tasks with high-level demands.

Work in the field has established that teachers who perceive students as struggling with concepts will respond by trying to build conceptual understanding (Copur-Gencturk et al., 2025). However, teachers' reported and actual instructional responses may differ, making it difficult to make inferences about the nature of their instructional responses to fractions misconceptions.

Further, prior work has shown that routinization and remediation are common ways the cognitive demand of high-level math tasks can decline during implementation, but these studies are smaller in scale and did not attend to the role of student misconceptions (Estrella et al., 2020; Henningsen & Stein, 1997). In one study that examined cognitive demand in teacher responses to student impasses during task implementation, Warshauer (2015) developed the "Productive Struggle Framework" to characterize the extent to which teachers keep students enlisted as problem solvers. The present study builds on this work by exploring the prevalence of similar teacher moves in a different content area, by examining how teachers support mathematical content connections of a task when responding to a student struggle, and by investigating how these content connections relate to pedagogical approaches that enlist the student as a problem solver. I argue that both mathematical connections and student accountability are dimensions of cognitive demand worthy of investigation because they constitute the "what" and "how" of task-based mathematics instruction (Doyle, 1988). For example, a teacher may enlist the student fully in the problem-solving of a task, but the "ask" may be reduced to completing algorithmic steps absent of connections. Conversely,

the teacher may take over the problem-solving of a task entirely while emphasizing to the student the connections between concepts, procedures, and meaning of the task's mathematics. Existing work on teacher preservation of task cognitive demand has primarily focused on student responsibility for problem solving, limiting our understanding of how teachers treat the meaning of the mathematics and of how they can preserve mathematical coherence while keeping the student enlisted. Moreover, studying task devolution and teacher responses to student challenges in situ is difficult due to myriad contextual and student-related factors (Warshauer, 2015); a methodology employing a problem-specific instructional vignette shows promise for characterizing the teacher action-related competence of preserving the cognitive demand of a high-level task disentangled from contextual factors that limit proximity, alignment, comparability, and occurrence of this core practice (Charalambous, 2020).

In the present study, I explore how teachers respond to a student struggling on a fractions task involving high-level cognitive processes. Specifically, I ask:

1. What do teachers focus on in terms of mathematical content about fractions, and to what extent does this content show connections between mathematical concepts, procedures, and underlying meanings intended by the task?
2. To what extent does the teacher hold the hypothetical student accountable for the high-level processes or products of the task?
3. How do dimensions of fractions content and its connectedness relate with student accountability for high-level processes and products in teachers' responses?

I carry out this study via an online survey that presented teachers with an instructional vignette about fractional equivalence. In the vignette, a hypothetical student shows their work on a mathematics task about finding out what remains of a pan of 16 brownies after $\frac{7}{8}$ of the brownies

have been sold. The student demonstrates a common misconception transferred from whole-number concepts, and the teacher is asked to record how they would respond to the student. The present investigation helps to illustrate how teachers preserve or diminish the high-level demands of a task to create—or hinder—student opportunities to learn fractions in K-8 mathematics, with implications for improving teacher knowledge, practice, and student learning in this crucial domain.

Literature Review

In this section, I provide an overview of the literature on connections between mathematical concepts, procedures, and their underlying meanings, arguing that these interconnections can address common student misconceptions and build enduring understandings of fractions. Throughout, I interweave examples from the literature on task-based fractions teaching and learning both with and without such connections to illustrate how connections can support a deeper learning of fractions. Next, I outline the importance of holding students accountable for high-level processes and products, i.e. *doing* math, to support enduring understandings of fractions.

Connecting Fraction Concepts, Procedures, and Their Underlying Meaning

Historically, conceptual knowledge in mathematics has been understood as “a connected web of knowledge” in which discrete facts and propositions are coherently organized in a network of linked relationships (Hiebert & Lefevre, 1986, pp. 3-4). Meanwhile, procedural knowledge was often reduced to “the ability to execute action sequences (i.e. procedures) to solve problems” (Rittle-Johnson & Schneider, 2015, p. 5) or “chains of prescriptions for manipulating symbols” (Hiebert & Lefevre, 1986, pp. 7-8). However, rather than viewing mathematical understanding along a conceptual-procedural binary, other scholars have pushed the field to instead view both procedural and conceptual mathematical as having the potential to be deep or superficial (Star, 2005). In other words, both mathematical procedures and concepts have the potential to be presented superficially

or deeply (Star, 2005). Deep mathematical knowledge—whether procedural or conceptual—is that which is interconnected into a coherent schema or network (Hiebert & Carpenter, 1992). Indeed, recent scholarship has emphasized the importance of intra-mathematical connections or linkages between mathematical ideas and objects (Murphy et al., 2024). When implementing high-level tasks, teachers play a key role in preserving cognitive demand through drawing “explicit connections between the mathematical ideas” (Henningsen & Stein, 1997, p. 526) and emphasizing the meaning of the mathematics involved (Doyle, 1988).

A core focus of reform in fractions teaching emphasizes the need to build meaning-oriented and interconnected conceptual and procedural understandings of fraction operations to help students understand *why* procedures make sense (Siegler et al., 2010). To illustrate, in their Mathematics Task Framework, Stein and Smith (1998) provide an example of a math task with high-level demands that focuses on connections between decimal, percent, and pictorial equivalents of the fraction $\frac{3}{5}$. Here, the student must work to represent the fraction $\frac{3}{5}$ on a **10x10** grid. When supporting students through such a task during implementation, the teacher can emphasize connections between concepts, procedures, and their underlying meaning— for example, by focusing on the meaning behind converting $\frac{3}{5}$ to the equivalent fraction $\frac{60}{100}$ — or diminish them, for instance, by solely focusing on the steps to multiplying the numerator and denominator of the fraction $\frac{3}{5}$ by 20. Indeed, a strict focus on procedural thinking is a common way students’ cognitive processes decline while working on high-level math tasks (Henningsen & Stein, 1997). As demonstrated by popular adages for fractions computations like “what you do to the bottom, you do to the top” for adding and subtracting, or “keep, change, flip” for dividing, common misconceptions about fractions operations have much to do with superficial understandings of

operations absent of meaningful connections (Fazio & Siegler, 2011; Fuchs et al., 2013; Siegler et al., 2010).

Other research demonstrates the importance of forging connections between concepts, procedures, and their meanings to address whole number misconceptions about fractions. Most student misconceptions about fractions can be traced back to not understanding fractions as numbers with magnitudes (e.g. Mack, 1993). Indeed, many U.S. students do not view fractions as rational numbers at all and view numerators and denominators as separate values, a misconception transferred from their understanding of whole numbers (Siegler et al., 2010). A “measurement understanding” of fractions, on the other hand, expands the number system and shows fractions as numbers with magnitudes that can be placed and “counted” on a number line, helping struggling learners build greater fluency with fractions (Fuchs et al., 2013; Siegler et al., 2010). Connecting fraction operations with a measurement understanding can give meaning to the values involved. For example, when supporting students through a task with high-level demands that starts with $\frac{9}{10}$ of a cake and asks student to determine how much would be left if $\frac{1}{2}$ were given away, the teacher could emphasize that $\frac{9}{10}$ is nearly one whole, and after subtracting half, roughly another half should remain. By forging connections between the fractional values and their magnitudes, students are supported in building a measurement understanding of fractions and a more coherent view of the rational number system beyond whole numbers (Behr et al., 1986; Siegler et al., 2010).

The use of visual representations to illustrate the meaning and conceptual basis behind common fraction operations may also support students in building an interconnected, coherent understanding of fractions (Siegler et al., 2010). In studies where fraction arithmetic is taught to students using visual representations, students’ computational skills and procedural fluency have shown improvement (Fazio & Siegler, 2011). Problem representation is thought to be a mechanism

for the links between the iterative process by which students build deep understanding in mathematics (Rittle-Johnson et al., 2001). Even for teachers solving word problems involving fractions, the use of pictorial representations is linked to successful problem solving (Copur-Gencturk & Doleck, 2021). Area models, diagrams, and other visual representations— such as fractions circles to demonstrate fractional equivalence and addition with common denominators (e.g. Cramer & Wyberg, 2009)— may help to demonstrate the interconnections between fractions concepts, procedures, and their underlying meanings.

Overall, whether teachers choose to focus on a measurement understanding of fractions or visual representations to convey their meaning, mathematics content must be presented as a coherent network of linked, interconnected relationships rich with meaning in order to resolve common misconceptions and support student mastery of fractions (Hiebert & Lefevre, 1986; Siegler et al., 2010).

Student Accountability for High-Level Processes and Products

Engagement in mathematics tasks with high levels of cognitive demand has been linked to student achievement gains in mathematics, for example, as seen in the QUASAR project (Silver & Stein 1996; Stein & Lane 1996). However, despite how a task may appear on paper, the task implementation stage is a critical phase of enactment during which the level of cognitive demand has the potential to devolve (Stein et al., 1996). Indeed, the TIMSS video project found that all high-level math tasks in their study devolved during instruction, underscoring the critical task implementation phase (National Center for Education Statistics [NCES], 2003; Stigler & Hiebert, 2004). As described above, high-level tasks can devolve into low-level tasks during implementation when the focus shifts to few or no connections between mathematical concepts, procedures, and their underlying meanings. Beyond reducing the interconnectedness of the mathematics, not holding

students accountable for high-level products or processes leads to the decline in the high-level cognitive demands of a task (Stein et al., 1996).

Doing mathematics necessitates engaging in the process of sensemaking (Schoenfeld, 1992). Holding students accountable for high-level products or processes in mathematics learning means enabling them to expend “effort to figure out something that is not immediately apparent” (Hiebert & Grouws, 2007). Much like building strength necessitates lifting weights in strength training, solving challenging mathematics tasks necessitates students grappling with and making sense of disciplinary ideas. Scholars have argued that student struggles or impasses are an important part of learning mathematics and something not to be overcome but rather used as an opportunity to guide student learning (Hiebert & Grouws, 2007; Warshawer, 2015).

Hurdles to keeping students enlisted as problem solvers during task-based math instruction include challenges with differentiating task support and addressing student push-back and struggles (Charalambous et al., 2023; Sullivan & Mornane, 2013). Teacher responses to student struggles vary according to the extent to which the problem-solving opportunity of the task remains intact for the student; in other words, for the student to still “do mathematics.” In response to a student struggling on a high-level math task, a teacher can provide the answer (Chazan & Ball, 1999), routinize or break the task down by directing the student to narrow features of the task (Estrella et al., 2020; Stein et al., 1996), provide step-by-step instructions or “recipes” to follow (Schoenfeld et al., 2023), probe or elicit from the student (Shaughnessy & Boerst, 2018), revoice the student’s thinking to reformulate or reframe the problem (O’Connor & Michaels, 1993), leverage student-generated mathematical representations in a process of guided co-construction (Terwel et al., 2009), have them retry the problem with minimal input, and more. Importantly, the ways in which teachers respond to students in these interactions determines how student agency and accountability are

distributed and thus how mathematical competence is constructed in the classroom (Gresalfi et al., 2009).

While the field lacks a comprehensive understanding of teacher responses to student struggles or stuck points in fractions teaching, emerging scholarship in this area demonstrates how teacher responses can resolve student challenges while preserving the cognitive demand of challenging tasks. In her exploratory case study of teacher responses to students struggles on high-level middle school math tasks about proportional reasoning, Warshauer (2015) offers a “Productive Struggle Framework” for characterizing teacher responses along a continuum from lower to higher cognitive demand. “Telling” responses are those in which the teacher supplies information or a strategy, evaluates student work, and corrects errors, leading to lowered cognitive demand of the enacted task. “Directed guidance” is that which redirects student thinking, narrows the scope of possibilities for action, directs some action, or breaks the problem down into smaller parts, which can serve to lower or maintain the cognitive demand of the task. “Probing guidance” occurs when the teacher asks for reasoning and justification from the student, maintaining the cognitive demand of the task as intended. Lastly, “affordances,” which can maintain or even raise cognitive demand of the task, are those in which the teacher builds on student ideas with limited but targeted intervention.

Relating Fractions Content and Student Accountability

Widely available guidance for effective fractions instruction centers primarily on the presentation of content, for example, the importance of fractions representations (Fazio & Siegler, 2011; Siegler et al., 2010). Moreover, the literature on cognitive demand tends to be domain-general within the discipline of mathematics, limiting an understanding of how it is instantiated within specific sub-domains in mathematics. While Warshauer’s (2015) examination of teacher responses to

student impasses during task implementation articulated concrete teacher actions in terms of how they enabled the student to grapple with the mathematics, the existing study builds on this work by also examining the mathematical meaning and connections in teachers' responses. For example, do certain practices, such as probing the student about the meaning of the denominator numerically and visually, work together to enlist the student as a problem solver *and* preserve mathematical meaning? An analysis that explores the relationship between teacher treatment of fractions content and level of student accountability would contribute to the field's understanding of cognitive demand within the specific and challenging sub-domain of fractions.

To summarize, teachers can preserve the cognitive demand of high-level tasks by emphasizing the connections between procedures, concepts, and their underlying meanings inherent to the task and keeping the student accountable for high-level mathematical processes and products (Boston, 2012; Hiebert & Grouws, 2007; Stein & Lane, 1996). As demonstrated by the literature in this area, the work of striking a balance between helping and challenging students who struggle when solving mathematics tasks is nuanced, raising the need for more scholarship on how teachers manage this tension (Stewart & Ball, 2023). To restate, in the present study, I investigate how teachers respond to a student struggling on a fractions task involving high-level cognitive processes. Specifically, I focus on two aspects of cognitive demand and their interrelations: how teachers represent the fractions content of the task and its connectedness, how the teacher holds the hypothetical student accountable for the high-level processes or products of the task, and how fractions content and its connectedness relates with student accountability for high-level processes and products.

Methods

In this section, I describe the study design and procedures, sample of teachers who participated, and approach to coding and analysis.

Study Design and Procedures

Teachers responded to an online survey administered via Qualtrics in the spring of 2025, which was the same survey as paper two in this dissertation. Participants were screened using two questions asking whether they have been a certified teacher of record in the U.S. with responsibility for instruction in a classroom that is between grades K and 8, and whether they teach mathematics as part of their teaching responsibilities. The survey informed participants they would be responding to teaching scenarios or “vignettes” that take place during student work time on math tasks, and that the scenarios featured the same 4th grade student and their work.

This study is based on the tradition that uses animated teaching simulations (ATSs), also called “vignettes,” to present teachers with an approximation of practice (Charalambous, 2020; Grossman et al., 2009; Herbst et al., 2011). A vignette is an instrument that offers “standardized yet realistic classroom prompts,” (p. 6) providing teachers with a narrowly defined performance opportunity at scale and in a standardized fashion (Hill et al., in press).

The survey presented teachers with an instructional vignette featuring a high-level fractions task and the work of a hypothetical 4th grade girl who has reached a stuck point, defined for participants simply as a place where the student has stopped making progress on the task. The task was created using free online resources by Illustrative Mathematics ® and the author’s own creation in Adobe Photoshop ©. The design of the math task and accompanying prompt featured in the vignette were further developed and finalized after holding cognitive interviews with five current or former elementary school teachers. The vignette described the topic, the math task with instructions,

the student's work in progress, and a description of what they had completed so far as well as their stuck point.

The mathematics of the task involved the concept of fractional equivalence: finding out what remains of a pan of 16 brownies after $\frac{7}{8}$ of the brownies were sold at a school bake sale. Teachers were presented with the student's work, which showed a 4x4 grid with seven squares shaded out of sixteen, demonstrating a common misconception transferred from whole-number concepts. Informed by Stein et al.'s (1996) task classification system, I interpret this task as one with high-level cognitive demands in that it involves procedures *with connections* and *doing* mathematics. In contrast to "familiar work," which relies on "memory or relatively standardized operations or algorithms," this task could also be interpreted as "novel work" because the student "must make decisions about what to produce and how to produce it" (Doyle, 1988, p. 173). For example, there is no standard solution method to this task, and the hypothetical student working on it created a representation and got stuck, showcasing the risk and high cognitive demands inherent to novel tasks.

Teacher responded to the prompt: "*What would you say to this student to support them in finding the number of brownies left?*" The prompt was intentionally neutral so that teachers could openly respond with a procedure, an explanation with mathematical connections, a probe to the student, and other approaches. Teachers responded by speaking out loud into their device's microphone. See Table 3.1 for a view of the fractions task and vignette with prompt. Teacher responses were recorded and transcribed using Voiceform software embedded into the survey. At the end of the survey, teachers responded to questions about their background.

Table 3.1. Vignette prompts and featured student work

Prompt	Mathematics Task and Student Work
Your class is working on the task below about equivalent fractions using a brownie pan cut into 16 pieces. You notice a student correctly represented 16 brownies, but they confused $\frac{7}{8}$ of 16 for $\frac{7}{16}$, as shown on their paper below. What would you say to this student to support them in finding the number of brownies left?	Lisa baked a large pan of brownies and cut them into 16 equal pieces. She sold $\frac{7}{8}$ of the brownies at the school bake sale. How many brownies does Lisa have left? 

Sample

The final sample included $N=228$ teachers who completed the survey and followed all survey prompts to record themselves speaking. The majority of teachers sampled were women (54.8%), yet women appeared in this data at a lower proportion compared to the national average in U.S. public schools (NCES, 2023). Teachers reported an average of 10.6 years of teaching experience. The level of instruction for 75% of teachers was between grades 2-6. Most teachers identified as White (59.2%), Black (19.3%), or Latinx/e (13.6%), which shows the sample had a more racially and ethnically diverse distribution compared to U.S. public schools (NCES, 2023). See Table 3.2 for a summary of the teacher sample.

Table 3.2. Sample Characteristics for Full Sample

	Sample (N=228)
	N (%)
Gender	
Woman	125 (54.8%)
Man	66 (28.9%)
Nonbinary	7 (3.1%)
Years of Experience	10.6 (8.3)
Grade level taught	
Kindergarten	8 (3.5%)
1 st grade	16 (7.0%)
2 nd grade	25 (11.0%)
3 rd grade	37 (16.2%)
4 th grade	44 (19.3%)
5 th grade	41 (18.0%)
6 th grade	23 (10.1%)
7 th grade	13 (5.7%)
8 th grade	20 (8.8%)
High School ⁷	1 (0.4%)
Teacher race-ethnicity	
Black	44 (19.3%)
Latinx/e	31 (13.6%)
Native Am.	2 (0.9%)
AAPI	6 (2.6%)
White	135 (59.2%)
Multiracial	10 (4.4%)

Note: AAPI stands for Asian American and Pacific Islander. Column percentages are shown in parentheses. Mean and SD are provided for years of teaching experience.

⁷ We retained $n=1$ high school teacher who met our screening criteria of having taught K-8 math in the past three years

Coding and Analysis

Transcribed responses were coded along two dimensions using a flexible coding approach: fraction concepts, procedures, and their underlying meaning (what I will call fractions content), and student accountability for high-level processes and products (what I will call student accountability). Both of these dimensions were informed by the literature on teacher implementation of math tasks with high-level demands (Boston, 2012; Stein & Lane, 1996). Flexible coding involved a combination of etic and emic coding, balancing staying low in the data while referring to existing theory to inform the coding scheme (Deterding & Waters, 2021; Miles et al., 2020). The unit of analysis was each teacher's full transcribed response to the prompt in the vignette. To code the data, I read through all responses inductively and recorded researcher memos of themes that emerged, staying close to the data to capture the fractions content and its connectedness and student accountability featured in the response without judgment (e.g. fractions content—response focuses solely on the steps in the standard algorithm for subtracting equivalent fractions; student accountability—the teacher dictates complete steps to solve the problem).

Next, I compared the memos to the literature on fractions instruction (Siegler et al., 2010) and teacher implementation of tasks with high-level demands, namely Warshawer's (2015) Productive Struggle Framework. From these comparisons, several themes emerged which became the codes for the codebooks for fractions content and its connectedness and student accountability, respectively. I consulted with two experts in mathematics education to complete multiple rounds of coding on randomly selected subsamples of the data with the developing codebooks. Through this iterative process of applying the two codebooks to the data, categorizing codes under themes, and re-applying the codebook to the data, I refined both codebooks. See Tables 3 and 4 for the codebooks for fractions content and its connectedness and student accountability, respectively.

Given that teachers often focused on numerous aspects of fractions content (e.g. understanding the numerator and denominator as values in relation, the standard algorithm for creating equivalent fractions), numerous fractions content codes could be applied to a response. Similarly, because teachers often exhibited multiple approaches to supporting the student in their responses (e.g. by first probing the student and then proceeding to outline steps to solve the problem), multiple student accountability codes could be applied to a response. In the findings section below, I describe the occurrences of individual codes in addition to drawing attention to how they occur in combination.

Two raters worked together to double-code 100% of the data for fractions content in batches, meeting in between sets to resolve discrepancies and reach agreement. Both raters were former mathematics educators and department chairs trained in qualitative research methods. They reached an average agreement rate of $Kappa=0.77$ across the fractions content codes. See Table A3.1 in Appendix A3 for a summary of the Kappa scores and positive agreement for math content. For the student accountability codes, the raters double coded 20% of responses in batches, reaching high levels of agreement early on and an average Kappa of 0.78. See Table A3.2 for a summary of the Kappa scores and positive agreement for student accountability.

Connecting Fraction Concepts, Procedures, and Their Underlying Meaning

The codes the above process produced can be categorized into the two areas of interest described above: the meaning oriented-ness of the teacher response, and whether students were held accountable for doing the work. In the first category, I grouped together codes according to the extent of the connections between fractions content reflected in them; in other words, the extent of mathematical meaning in the response as informed by the literature on effective fractions instruction (Fazio & Siegler, 2011; Siegler et al., 2010) and based on the definition of intramathematical

connections as linkages between mathematical objects such as “ideas, concepts, theorems, procedures, or representations within mathematics” (Murphy et al., 2024, p. 231). For example, the code “connects the visual representation to the problem context” was applied to responses where the teacher drew connections between the visual fractions diagram and the fractional values in the problem, e.g. “16 brownie squares tells us the denominator of our fraction is sixteenths.” Here, the teacher connects the denominator to the whole of the fraction in the diagram. Other fractions content codes with connections were focused on the standard algorithm for creating equivalent fractions with connection to the meaning behind the steps, understanding the numerator and denominator as values in relation, and connecting the fraction $\frac{7}{8}$ to its measurement value or magnitude, e.g. nearly one whole.

Two individual codes described responses with no connections between fraction concepts, procedures, and their underlying meaning— those that centered on the standard algorithm for creating equivalent fractions focused solely on operations, and those that stated $\frac{7}{8}$ and $\frac{7}{16}$ are not the same value without any elaboration. The latter code was a simple correction without any explanation of why the two values differ; previous scholarship has shown a singular focus on correctness lowers the cognitive demand of high-level tasks (Estrella et al., 2020).

Table 3.3. Codebook for Fractions Content and Its Connectedness

Connectedness	Fractions Content Code	Criteria	Example
<i>Connects Fraction Concepts, Procedures, and/or Their Underlying Meaning</i>	Standard algorithm for creating equivalent fractions focused on the meaning behind the steps	Discusses the concepts and meaning underpinning the algorithm, e.g. why it makes sense	“We need to divide the 16 brownies into 8 equal groups. How many brownies would be in each group, right? Each 8 would have 2 brownies because 16 divided by 8 is 2. Now, if you have 7 out of those 8 groups, how many brownies is that? Yes, it’s 7 times 2, which equals 14. So, seven-eighths of 16 is 14 brownies, not seven-sixteenths.”
	Understanding the numerator and denominator as values in relation	The response focuses on the interpreting the numerator in relation to the denominator, e.g. as a rate	“It’s not just that they sold 7 brownies but 7 out of 8, which is almost one whole! You have to look at both parts, not just the top number.”
	Connects the visual representation to the problem context	The response links aspects of the problem to the visual diagram, e.g. by mentioning “shading” or “boxes” or “squares”	“I see you drew 16 pieces to represent the 16 brownies. Can you show me where you see $\frac{7}{8}$ in the boxes you drew?” “How can we break up these pieces into eight parts?”
	Estimating/ intuiting the value of $\frac{7}{8}$, i.e. that it is nearly one whole	Focuses on the magnitude or value of the fractions, e.g. by referencing landmark fractions like $\frac{1}{2}$ or one whole	“Seven out of eight is almost eight out of eight, which would be one whole. They sold almost all the brownies.”
<i>Does Not Connect Fraction Concepts, Procedures, or Their Underlying Meaning</i>	Standard algorithm for creating equivalent fractions focused on operations	Recitation of steps to the operations without connection to meaning. Can be specific or general	“First you need to convert seven-eighths into fourteen-sixteenths. You do that by multiplying the numerator and denominator by two. Then, you can subtract!” “Whatever you do to the bottom, you have to do to the top. Remember how to make an equivalent fraction?”
	The fact that $\frac{7}{8}$ is not the same as $\frac{7}{16}$ without elaboration	Simple statement of fact that the two values are not the same	“You drew seven-sixteenths instead of seven-eighths. Those are not the same”

Table 3.3. Continued

<i>Other (fractions content and connections undetermined)</i>	No math content	The response does not mention any mathematics, whether vague or specific	“Meet me after class”
	Mathematical errors and/or imprecision	The response contains mathematical errors, imprecise use of language, or lack of clarity which makes it confusing	“When you subtract, your answer will be nine-sixteenths.”
	Unspecified mathematical focus	The math content is vague and unspecific	“Let’s practice a different, easier problem”

Note: Multiple fractions content codes could be applied to a teacher response.

In responses with an unspecified or lack of mathematics content, for which extent of mathematical connections was indeterminable, the code “unspecified math focus” or “no math content” was applied, respectively. For teachers who exclusively elected to probe the hypothetical student (e.g. “How did you come up with that?”) or direct them to retry the problem (e.g. “I wonder if you can go back and find your mistake and try again”), the mathematics content is not explicitly ascertainable due to the nature of these pedagogical approaches. To distinguish between these types of responses and those in which the teacher ostensibly did not choose to engage fractions content, I summarize fractions content and its connections exclusive of responses that did not contain explicit mathematics content due to the teacher electing to pose a question or direct a student to retry. In the limitations section, I revisit this issue more broadly as it relates to the vignette methodology.

Finally, responses that exhibited mathematical errors, imprecision, or were mathematically confusing were tagged with the code “mathematical errors and/or imprecision.” In the findings, I describe trends in responses coded for math errors. See Table 3.3 for a summary of the fractions content codes and the connections represented in them.

Next, I organize the math content findings according to the extent of the content codes *with connections* between fraction concepts, procedures, and their meaning applied to the responses (in

addition to any code from other categories). Specifically, I start with an overview of responses that showed multiple connections—those which contained at least two individual codes with connections. I follow with those that showed some connection, which had a single individual code with connections applied, and conclude with a discussion of responses which were not coded for any connections. Importantly, this section excludes the 16 responses noted earlier which did not exhibit surmisable mathematics content due to the pedagogical approaches of probing or directing the student to retry. Organizing the codes at this level enables an analysis that goes beyond individual content codes to illustrate more broadly how meaning-dense a teacher’s response was.

Student Accountability for High-Level Processes and Products

In the second area of interest, I organize the codes produced by the process above according to the nature of student accountability for problem solving. Informed by Warshauer’s (2015) productive struggle framework, I grouped together codes according to the extent to which the teacher takes over solving the problem of the task. Warshauer (2015) outlines four teacher approaches to supporting students through an impasse on a high-level mathematics task, resulting in varied levels of task cognitive demand: telling, directed guidance, probing, and affordances. Informed by this framework and applying it to the standardized setting of the vignette absent of the student, I similarly qualified teacher responses according to who was responsible for solving the problem (the teacher or the hypothetical student). See Table 3.4 for a summary of student accountability codes.

Probing involved asking or pressing for reasoning and justification, asking questions about the student’s approach, or asking the student about the reasonableness of their answer. Responses in which teachers probed the student (e.g. “What does the fraction $\frac{7}{8}$ mean? Can you show me where your denominator shows up on the diagram?”) were considered higher in their preservation of task

cognitive demand because the questions were open-ended, and the problem-solving opportunity remained fully intact for the hypothetical student. Directing the student to retry involved the teacher telling the hypothetical student to try the problem again. This pedagogical approach took the form of restating or having the student restate the question, or generally telling the student to find their mistake, for example, by stating, “Go back and re-read the question and make sure you’ve marked your representation correctly” or “Let’s correct our mistakes.” These responses also preserved the problem-solving opportunity for the student by prompting them to diagnose their own errors and try the problem again.

The code “scaffolding” was applied to responses in which the problem solving was shared between the teacher and hypothetical student, leaving space for the student to think, reason, or grapple with the problem. Scaffolding involved the teacher providing a hint, partial solution or explanation, or enabling prompt to address the student’s misconception and reduce problem complexity without proposing a specific solution pathway (as in telling below) (Anderson, 1989; Henningsen & Stein, 1997; Sullivan et al., 2006). This could be through a statement or an explanation, with the problem-solving aspect of the task partially intact for the student to carry out, or through a back-and-forth exchange similar to an initiation-response-evaluation (I-R-E) sequence (Cazden, 1988). Importantly, questions coded for a pedagogy of scaffolding differed from probing questions in that they were not open-ended questions.

Table 3.4. Codebook for Student Accountability for High-Level Processes and Products

<i>Student Accountability</i>	Code	Criteria	Example
<i>High</i>	Probing	The teacher poses open-ended questions to the student asking for reasoning and justification, asking about the student's approach, and asking the student about the reasonableness of their answer.	"See this number seven-eighths, what does that mean?" "Can you explain what the meaning of the denominator is?"
<i>High</i>	Direction to retry	The teacher directs the student to try the problem again, either by reading the question for them, having the student re-read the question, or having the student diagnose their own error.	"Alright kiddo, I want you to double check your work because I think you have something to fix here."
<i>Moderate</i>	Scaffolding	The teacher provides a hint, partial solution or explanation, or enabling prompt to address the student's misconception and reduce problem complexity without proposing a specific solution pathway (as in telling). This could be through a statement or an explanation, with the problem-solving aspect partially intact for the student to carry out, or the back-and-forth exchange of funneling questions (e.g. I-R-E included in sequence) so that the problem solving is shared.	"Hm ok let's think about the denominator of this problem. The brownie has 16 squares, and we want seven-eighths of it—not the number 7, but the amount $\frac{7}{8}$ represents."
<i>Moderate</i>	Privileging student strategy	The teacher references, revoices, restates, or summarizes what the student has already done and builds on the student's initial thinking. The teacher can also leverage the student's initial strategy in their response.	"So I see you colored in seven out of eight but then there's a whole nother section of eight there that you didn't color in. So really if you look at that pan of brownies the way you colored in, you colored in seven sixteenths. Is seven-sixteenths equal to seven-eighths?" *
<i>Low</i>	Telling	The teacher provides a complete or near-complete process for solving the problem, i.e. a teacher-led or teacher-proposed solution pathway. This also includes evaluations of correctness by the teacher.	"You need to convert seven-eighths into an equivalent fraction. Multiply the top and bottom by 2."
<i>N/A</i>	Unspecified pedagogical approach	The response is vague or non-specific in its pedagogical approach.	"Let's rewrite this out on the board together."

Table 3.4. Continued

Note: Multiple student accountability codes could be applied to a teacher response. * This response was also coded as scaffolding.

When the teacher referenced, restated, revoiced, summarized or leveraged the hypothetical student's initial strategy to build on their thinking, the code "privileging the student strategy" was applied. Because the initial student strategy was to draw seven shaded boxes out of sixteen, I applied this code to responses that leveraged and built on this representation in any way. Following work by Terwel et al. (2009), this approach characterizes a scaffolding process of "guided co-construction" in which the teacher builds on a student's mathematical representation or model, which may support students in developing "a sense of knowledge ownership" (p. 41) and transfer of learning. Due to the problem-solving aspect of the task being shared between the teacher and hypothetical student, and thus accountability and agency being distributed (Gresalfi et al., 2009), both scaffolding and privileging the student strategy were qualified as showing moderate student accountability and thus partial preservation of task cognitive demand.

The code "telling" was applied to responses in which the teacher provided a complete or near-complete process for solving the problem. Responses of this nature involved the teacher taking over the problem-solving process for the student, either by fully articulating or proposing a solution strategy. Because the hypothetical student was no longer responsible for any high-level process or product, this type of response demonstrated devolution of task cognitive demand. Finally, responses with a vague or unspecified pedagogical approach were coded as such.

Relating Fractions Content and Student Accountability

The final strand of inquiry is concerned with how teacher handling of student accountability and fractions content relate with one another. I first cross-tabulate individual fractions content codes by the student accountability codes to demonstrate their co-occurrences. Next, I take a

broader view to show the co-occurrences between these same student accountability codes, this time across varying levels of mathematical connections (no, some, or multiple). For these analyses, I focus on responses that had codable fractions content and pedagogical approaches for student accountability. I present findings descriptively and with 100% stacked bar charts to demonstrate the proportion of content and accountability codes that co-occurred with one another. This high-level view showing how specific pedagogical moves relate to the meaning-density of responses enables us to understand how teachers manage keeping the student enlisted while preserving the mathematical coherence of high-level tasks.

Findings

I present findings in terms of the connectedness of fractions content and student accountability for solving conveyed in teachers' responses. First, I describe the raw frequencies and proportions of individual fraction content codes. Because teachers often discussed various fractions ideas (and multiple codes could apply), I follow by drawing attention to how teachers engaged with fractions content to forge multiple, some, or no connections between fraction concepts, procedures, and their meaning. Next, I discuss the extent to which teachers held the hypothetical student accountable for high-level processes and products. Findings for fractions content and student accountability are presented separately; however, when appropriate, I describe the ways in which content connections and degree of student accountability work together within a response to maintain (or diminish) the high-level demands of the task. I dedicate the final section of the findings to outlining patterns in the connections between these two aspects.

Connecting Fraction Concepts, Procedures, and Their Underlying Meaning

I begin with a discussion of the fractions content and its connectedness agnostic of student accountability, which I cover separately in the second section of the findings.

Individual Fractions Content Codes

In a subset of responses (12.3%), teachers forged connections by underscoring the concepts and meaning behind the standard algorithm for finding equivalent fractions. Nearly a third of teachers (32.0%) drew connections between the visual representation and the problem context. Few responses (8.3%) discussed the numerator and denominator as values in relation, and even fewer (4.4%) were concerned with estimating the measurement value or magnitude of $\frac{7}{8}$. See Table 3.5 for a summary of the frequencies and proportions for the individual fractions content codes.

Nearly a quarter of responses (24.1%) focused on the operational steps behind the standard algorithm for creating and subtracting equivalent fractions, e.g. “Seven-eighths needs to be timesed by two, so you get sixteen equal squares, and you’d have fourteen out of sixteen. So that would leave two brownies.” Of these responses focused on standard algorithmic steps, only 12.7% showed connections between the visual and problem context; meanwhile, 30.9% of these also focused on the fact that $\frac{7}{8}$ and $\frac{7}{16}$ were not the same value without elaboration, and 40% of these contained mathematical errors and/or imprecision.

In over one in five responses (21.5%), teachers focused on the simple fact that $\frac{7}{8}$ and $\frac{7}{16}$ were not the same value numerically. These responses were simple corrections and stopped short of any meaning-oriented discussion of the difference between the two fractions. For example, one teacher responded: “... So for that one, you’re using seven of the sixteen pieces, but you gotta realize that they’re... We’re talking about seven eighths...” In a small portion of these responses (14.3%), teachers also focused on the meaning behind standard algorithm. In an even smaller portion of these responses (12.2%), the teacher also discussed the numerator and denominator as values in relation.

A small portion of responses (4.4%) contained no surmisable mathematics content, general or otherwise (e.g. “you have misunderstood the instructions”), while others (25.9%) were nonspecific in their mathematics focus. In 16 of these nonspecific responses, the teacher did not explicitly specify fractions content due to taking a pedagogical approach of exclusively probing the student (e.g. “Now let’s see, if these were your brownies, how would you like it to be done?”) or directing them to retry (e.g. “Please review your answer again”). In the former, the teacher appears to try and probe the student to see a greater portion of brownies sold. Nonspecific responses suggested some mathematical activity but were generic and somewhat cursory. For example, one teacher shared, “Okay, Student, let me help you with this problem. Let’s go back and draw this problem according to all the information that we have, not just some of the information that we have.” Here, the teacher mentions “drawing the problem” according to more information, but it is unclear what the mathematical focus is. In a portion of all responses with an unspecified mathematical focus (16.9%), the teacher outlined an unspecified alternative task or prior activity, for example an easier, similar problem or hypothetical lessons practiced previously.

Throughout all responses with codable mathematics content, 26.8 % contained some sort of mathematical error (e.g. “nine brownies would be left”) or were confusing due to a lack of precision in language. The most common types of errors involved the teacher saying the student was correct, that there would be nine brownies remaining (e.g. “count the remaining squares to find your answer”), or that the student only needed to show eight brownies in their diagram instead of sixteen (e.g. “make a smaller diagram”).

Table 3.5. Frequencies and Proportions for Individual Fractions Content Codes

	Fractions Content Code	Frequency	Proportion
<i>Connects Fraction Concepts, Procedures, and/or Their Underlying Meaning</i>	Standard algorithm for creating equivalent fractions focused on the meaning behind the steps	28	12.3%
	Understanding the numerator and denominator as values in relation	19	8.3%
	Connects the visual representation to the problem context	73	32.0%
	Estimating/ intuiting the value of $\frac{7}{8}$, i.e. that it is nearly one whole	10	4.4%
<i>Does Not Connect Fraction Concepts, Procedures, or Their Underlying Meaning</i>	Standard algorithm for creating equivalent fractions focused on operations	55	24.1%
	The fact that $\frac{7}{8}$ is not the same as $\frac{7}{16}$ without elaboration	49	21.5%
<i>Other (fractions content and connections undetermined)</i>	No math content	10	4.4%
	Mathematical errors and/or imprecision	61	26.8%
	Unspecified mathematical focus	43	18.9%
	Unspecified mathematical focus due to pedagogical approach	16	7.0%

Note: Multiple fractions content codes could be applied to a teacher response.

Multiple connections between fraction concepts, procedures, and their meaning

Teachers who drew multiple connections between fraction concepts, procedures, and/or their underlying meaning preserved the meaning of the mathematics of the task. However, this occurred in only 10.8% of responses. To illustrate, one teacher stated:

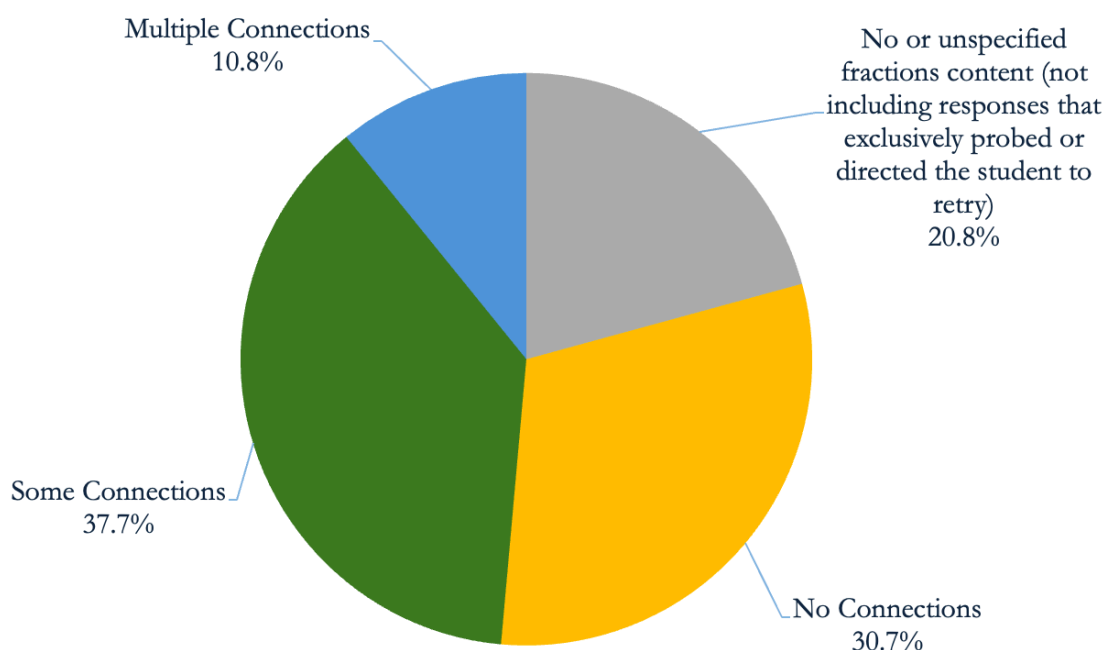
Okay, that was a great job you did, representing the whole brownie pan with sixteen pieces. Now let's clarify the difference between seven-eighths and seven-sixteenths. Remember the

bottom number, the denominator right, tells us how many equal parts the whole is divided into. So seven-eighths means we're looking for seven out of eight equal parts of the whole pan. Seven-sixteenths on the other hand would mean we're looking for at seven out of sixteen equal parts. To find seven-eighths of sixteen we need to think about dividing the pan into eight equal sections first. How many brownies would be in each of those eight sections? Once you know that you can figure out how many brownies would be in seven of those sections a piece, which would be seven-eighths of the whole pan. Think about it this way, seven-eighths is a large portion of the whole, now seven-sixteenths is just a little bit, so you're looking for almost the entire pan, not just a small part of it.

Here, the teacher explains the difference between seven-eighths and seven-sixteenths by showing the numerator and denominator as being in relation (e.g. "we're looking for seven out of eight equal parts of the whole, not seven out of sixteen equal parts"). She also forges connections between the number of equal parts the pan is divided into and the denominator of the fraction, making notable connections between the problem context and visual representation. This response also features relevant mathematical vocabulary, with mention of "equal parts" and the meaning of the denominator. Finally, the teacher emphasizes the measurement value and magnitude of seven-eighths, stating it is "almost the entire pan, not just a small part of it." See Figure 3.1 for a summary of the proportion of responses that featured multiple, some, none, or an undetermined number of connections between fraction concepts, procedures, and their meaning.

Responses that drew multiple connections most often did so by leveraging the visual representation. For example, in a large portion (39.3%) of responses which focused on the meaning behind the standard algorithm also forged connections between the visual and problem context, and nearly half (47.4%) of responses that related the numerator and denominator also drew connections between the visual and problem context. These co-occurrences indicate that the visual representation plays a pronounced role in making meaning of fractions.

Figure 3.1. Connections between fraction concepts, procedures, and their meaning ($n=212$)



Note: $n=212$ here due to excluding the $n=16$ which did not explicitly specify fractions content due to the teacher electing to probe or direct the student to retry.

Some connections between fraction concepts, procedures, and their meaning

A number of responses (37.7%) exhibited some connections between fraction concepts, procedures, and their meaning. For example, one teacher stated, "...it is seven-eighths of sixteen, not seven out of sixteen. Seven-eighths of sixteen would be higher than seven-sixteenths, which wouldn't even be half." In this response, the teacher briefly notes that seven-eighths of sixteen and seven-sixteenths are not the same value without much elaboration. Then, they leverage a measurement understanding of both fractions by comparing their magnitudes. This teacher uses landmark fractional values that are most familiar to students (e.g. one half, one whole) to provide a reference point for determining the relative value of fractions.

Most responses achieved some level of connection between fraction concepts, procedures, and their meaning by connecting the visual representation and problem context (60.0% of the “some connections” category). For example, one teacher stated:

Student, you’re right that they sold seven of the brownies, but look at your pan of brownies again, and notice there’s not eight pieces there, there’s sixteen pieces there. So, I would divide that pan of brownies straight across in half, and show that they have seven out of eight in the top half, then do seven out of eight in the bottom half, to see how many brownies she actually has left.

Here, the teacher draws connections between the denominator and the total number of parts in the brownie pan. They leverage the student’s diagram to create eighths visually, bypassing the need to create an equivalent fraction by other means.

No connections between fraction concepts, procedures, and their meaning

In 30.7% of responses, teachers made no notable connections between fraction concepts, procedures, and their underlying meaning even though they discussed the mathematics content of the task. The most common way teachers did this was by focusing on the steps to the standard algorithm for creating and subtracting equivalent fractions (70.8 %). For example, one teacher responded, “So you first need to find a common denominator for seven-eighths and sixteen and the common denominator would be two, so eight times two is sixteen and seven times two is fourteen. That would mean that fourteen of the sixteen brownies were sold.” Here, the response solely focuses on operational steps. Moreover, there is an error when the teacher says the common denominator is two rather than sixteen. Other responses were similar and focused solely on the standard algorithm for creating equivalent fractions more generally, stating “When simplifying fractions, we always need to make sure that we simplify not only the top number, but also the bottom number and vice versa.”

Lastly, 20.8% of responses did not specify any mathematics or fractions content, making the extent of connections indeterminable. This set of responses excludes the path dependency stemming from responses that exclusively probed or directed the student to retry. See the section above on responses with a lack of or nonspecific fractions content. Overall, I find that nearly two-thirds of responses coded for fractions content made at least one connection between fractions concepts, procedures, and their underlying meaning. A majority of these responses forged a connection by connecting the problem context and the visual fraction diagram. Few teachers forged multiple connections between fraction concepts, procedures, and their underlying meaning, and a large portion of teachers did not forge any connections even though they attended to the mathematics of the task. When no connections were made in teachers' treatment of fractions content, it was due to a singular focus on the steps in the standard algorithm for creating and subtracting equivalent fractions.

Student Accountability for High-Level Processes and Products

In this section, I discuss the pedagogical approaches teachers used to hold the student accountable for high-level processes or products. Responses ranged according to how the teacher kept the hypothetical student enlisted as a problem solver. Here, I outline the findings into pedagogical themes of probing, directing the student to retry, scaffolding, privileging the student strategy, and telling. See Table 3.6 for a summary of the frequency and proportion of student accountability codes applied to responses.

Probing

A small proportion of responses featured a pedagogy of probing (7.5%). To illustrate, one teacher responded with, "Let's read through the question carefully again. See this number seven-eighths, what does that mean?" Here, the teacher open-endedly asks the hypothetical student to

share their reasoning about the meaning of the fraction seven-eighths. However rare probing was, it most often occurred in combination with other pedagogical approaches. Interestingly, 41.2% of responses in which the teacher probed the student also privileged the student strategy. One teacher responded, “Great drawing of the 16 boxes. How’d you come up with that?” Here, the teacher leverages the student’s model and proceeds to ask them to share the mathematical reasoning behind their approach. In 29.4% of responses in which the teacher probed the student, the teacher did so in combination with scaffolding. For example, one teacher stated:” OK... your model is off to a great start. I see you made 16 equal parts just like you were supposed to. Then you shaded 7 parts. What would that fraction be? $\frac{7}{8}$ or $\frac{7}{16}$? How could we make it be eighths and not sixteenths? Here, the teacher scaffolds by targeting the misconception and asking the student to name the fraction in their diagram, following an I-R-E pattern of questioning. At the end, they probe the student, asking them how they might make eighths out of sixteenths in an open-ended fashion, inviting them to share their reasoning. The teacher also privileges the student’s strategy here, leveraging their diagram to distinguish between the fractions $\frac{7}{8}$ and $\frac{7}{16}$ and target the misconception.

Table 3.6. Frequencies and Proportions for Student Accountability Codes

Student accountability for high-level process or product	Pedagogy Code	Frequency	Proportion
<i>High</i>	Probing	17	7.5%
<i>High</i>	Direction to retry	41	18.0%
<i>Moderate</i>	Scaffolding	67	29.4%
<i>Moderate</i>	Privileging student strategy	45	19.7%
<i>Low</i>	Telling	119	52.2%
<i>N/A</i>	Unspecified pedagogical approach	40	17.5%

Note: Multiple student accountability codes could be applied to a teacher response.

Direction to re-solve problem

In 18.0% of responses, the teacher directed the student to try the problem again. This pedagogical approach took the form of restating or having the student restate the question, or generally telling the student to find their mistake, for example, by stating, “Go back and re-read the question and make sure you’ve marked your representation correctly” or “Let’s correct our mistakes.” These responses entirely preserved the problem-solving opportunity for the student by prompting them to diagnose their own errors and try the problem again.

Scaffolding

The second most common pedagogical approach, observed in 29.4% of responses, was scaffolding. For example, one teacher responded, “Okay, can we look at the denominator we are supposed to be using? We have sixteen squares, and we are looking for seven-eighths of the brownie, so we’re not looking for the number, we are looking for how much seven-eighths is.” In this response, the teacher reduces the task’s complexity by focusing the student’s attention on the denominator, addressing the student’s misconception while keeping the problem-solving aspect intact for the student. The math content featured here relates the numerator and denominator as a whole rather than separate values (“we’re not looking for the number [seven]”) and connects the diagram and the problem context by relating the sixteen pieces to the denominator.

Nearly a third (29.4%) of responses coded for a pedagogy of scaffolding also privileged the student strategy. For example, one teacher said,

Go back and look at your picture and see, look at the whole pan, and see how many pieces are there in total, which is going to be one of the main parts of our fraction. So you have seven pieces that are still there, but how many pieces were there in all?

Here, the teacher leverages the student's diagram to relate the concept of denominator to the total number of equal parts in the pan. This scaffold that leverages the student's initial strategy helps to meaningfully differentiate between the fractions $\frac{7}{8}$ and $\frac{7}{16}$.

Importantly, over one in five responses coded for scaffolding (22.1%) also included a pedagogy of telling. For example, one teacher responded:

Okay, in order to figure this one out, you've got to make it that seven-eighths into the equivalent to 16 equal pieces. So you have to figure out seven-eighths of 16. So you've got to make sure that that bottom number of the seven-eighths is 16. So you would do that and you would find out that it's fourteen-sixteenths. So you would say fourteen-sixteenths minus fourteen-sixteenths. So you would have two-sixteenths.

In the first two sentences, the teacher addresses the misconception without prescribing a specific solution pathway. In other words, the student could choose their own strategy for determining how to partition sixteen pieces into eighths, even though the teacher has reduced the complexity of the problem. However, the response devolves into telling, in which the teacher specifies the standard algorithm and even goes on to solve the entire problem themselves. Lastly, the teacher misspeaks towards the end when they mean to say we subtract 14 from 16.

Privileging the student strategy

Nearly one in five teachers (19.7%) privileged the hypothetical student's strategy in their response. Some leveraged the student's initial mathematical ideas, while others revoiced or summarized what the student had already done. When privileging the student strategy, the teacher often did so in combination with a pedagogy of scaffolding or probing, as described above.

Telling

A majority of responses (52.2%) demonstrated a pedagogical approach of telling. Responses of this nature involved the teacher taking over the problem-solving process for the student, either by

fully articulating or proposing a solution strategy. An example of this approach is shown in the following response:

Alright, you look at the fraction, seven-eighths. Well, we have to really figure out seven-sixteenths. So what do we do here to make this fraction bigger than seven-eighths? Well, we have to multiply each of these numbers by two. So you multiply the eight by two and the seven by two. So you say, oh, it really is fourteen- sixteenths.

In this example, the teacher carries out the cognitive work of creating an equivalent fraction and supplies the student with the necessary computations. Additionally, this response demonstrates some confusion or imprecision around fraction magnitude (that $\frac{14}{16}$ is “bigger” rather than equivalent to $\frac{7}{8}$). Notably, only 16.7% of responses coded for a pedagogy of telling also privileged the student’s strategy, underscoring the one-sided nature of problem solving that happens with telling.

Unspecified pedagogical approach

In a portion of responses (17.5%), the teacher’s response was too vague and nonspecific to ascertain their pedagogical approach, e.g. “Why don’t you grab another paper so we can correctly identify the fractions and I will help you.” Here, the nature of the help is unclear, as the response does not specify how the teacher would support the student in identifying the fractions.

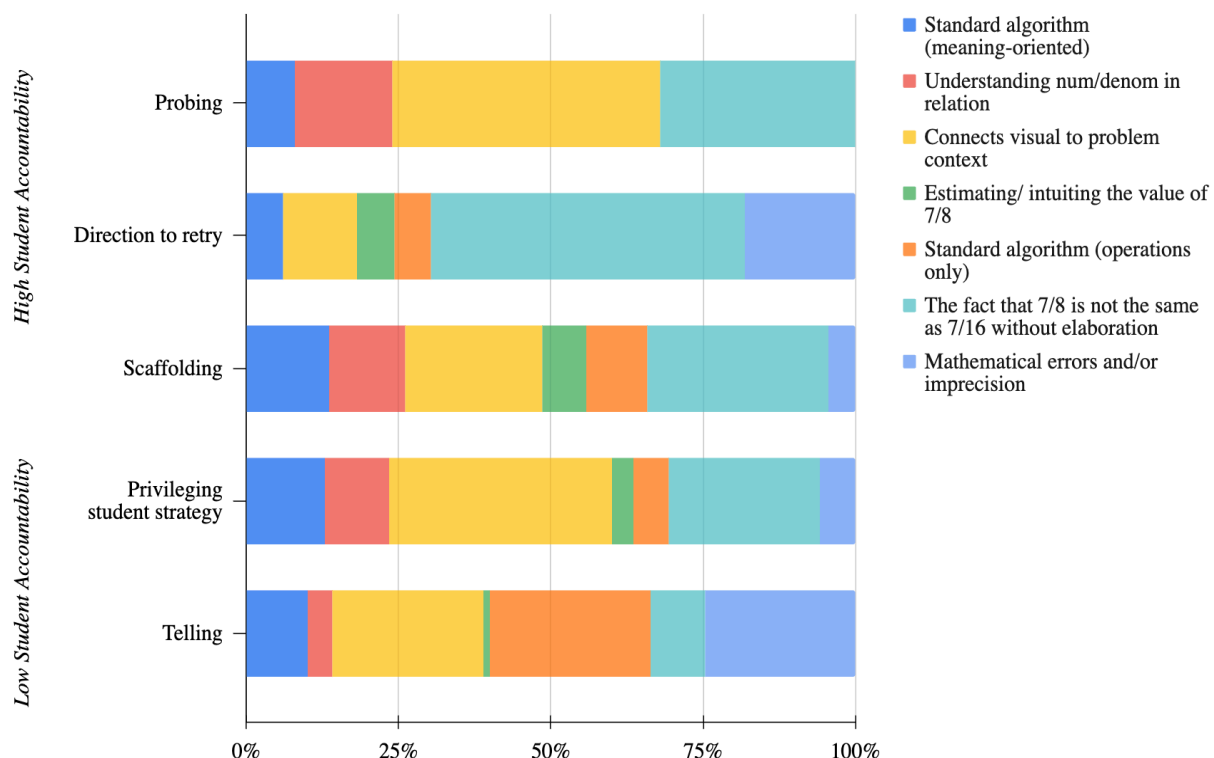
In sum, I find that probing the student was rare, but it was fairly common for the teacher to direct the student to try the problem again. In about a third of responses, teachers took the approach of scaffolding, reducing problem complexity while leaving the problem open for the student to solve and thus preserving some of the task’s cognitive demand. Nearly one in five teachers privileged the student strategy in some way. Interestingly, scaffolding and probing often co-occurred with privileging the student strategy, suggesting a potentially complimentary role between these pedagogical approaches. Lastly, a majority of pedagogical approaches—over half—involved teacher-directed problem solving, resulting in the hypothetical student not being accountable for any

high-level mathematical processes or products and thus reducing the high-level demands of the task as intended.

Relating Fractions Content and Student Accountability

Looking at patterns in the co-occurrences between individual fractions content codes and student accountability codes, I find that a focus on standard algorithmic steps most often occurred via a pedagogy of telling. That is, respondents who took over the problem solving most often did so via the steps to the standard algorithm. Moreover, mathematical errors or imprecision featured prominently with a pedagogy of telling, likely due to a focus on step-by-step procedures. When the teacher took a scaffolding approach or privileged the student strategy, a greater variety of fractions content foci featured in their responses. While probing was rare, it never featured a focus on the steps to the standard algorithm for creating equivalent fractions or estimating the measurement value of fractions; instead, probes heavily leveraged the visual diagram presented in the student's work. Connecting the visual fractions representation to the problem context was common in all pedagogical approaches but less so in directions to retry. Directions to re-solve the problem most often engaged simple corrections around the fact that $\frac{7}{8}$ and $\frac{7}{16}$ are not the same without elaboration. See Figure 3.2 for a summary. While descriptive, these findings suggest potential synergies between mathematics content fluency in the domain of fractions and pedagogical skill of enlisting the student.

Figure 3.2. Individual Fractions Content Codes by Student Accountability Codes

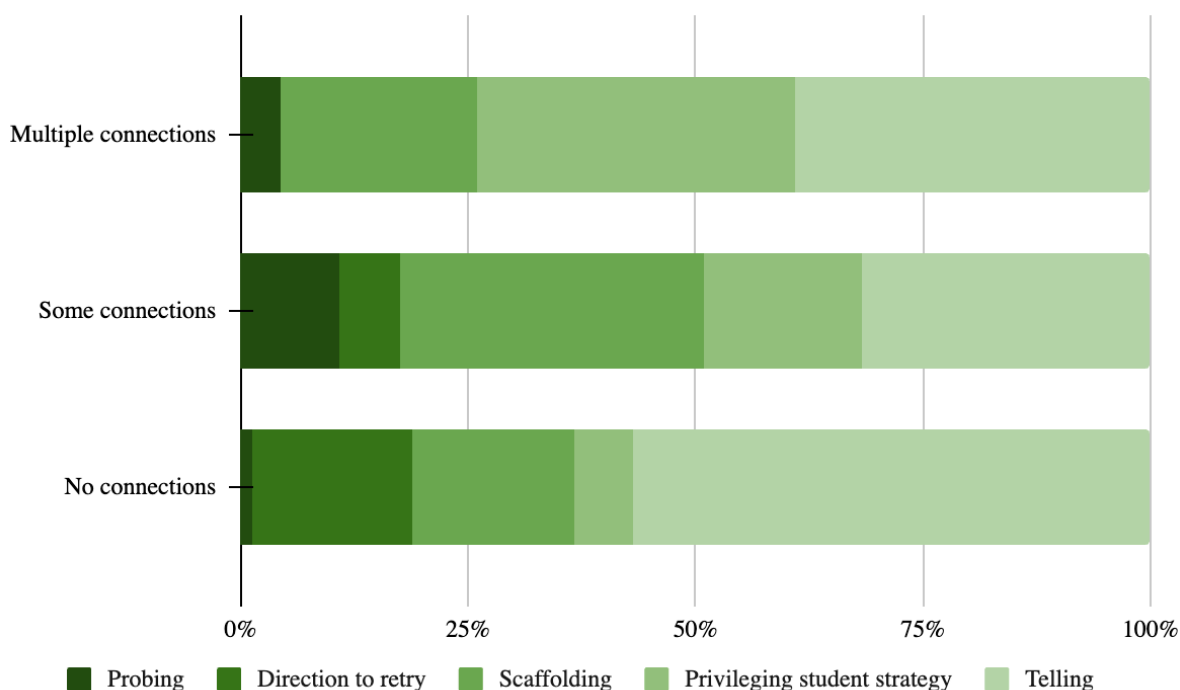


Note: In this bar chart, I focus on the cross-tabulations between responses that had codable fractions content and pedagogical approaches for student accountability.

Looking more broadly at the relationship between mathematical connections and student accountability in teacher responses, some notable patterns emerge expanding on the findings above. While telling approaches were common across the board, telling appeared most prominently in responses that did not forge connections between fraction concepts, procedures, and their underlying meaning. On the other hand, responses where teachers privileged the student's strategy most often drew multiple connections. Scaffolding approaches emerged similarly across responses with all levels of connection, but they were most common in responses that forged some level of connection. Approaches that directed the student to retry rarely appeared in responses that drew multiple connections and indeed often appeared without connections, which is expected from this type of pedagogical approach that puts the problem entirely back on the student with limited teacher intervention. Lastly, while it was rare for the teacher to probe, responses in which teachers did tend

to feature at least some level of connection between fraction concepts, procedures, and their underlying meaning. See Figure 3.3 for a summary.

Figure 3.3. Relating Mathematical Connections and Student Accountability



Note: In this bar chart, I focus on the cross-tabulations between responses that had codable fractions content and pedagogical approaches for student accountability.

To conclude, I discuss notable patterns in the relationship between mathematical connections and student accountability. First, while a high level of student accountability was rare in teacher responses overall, when teachers did enlist the student as a problem solver—whether by scaffolding or by privileging their strategy—it was more likely to occur in responses that drew some connections between fraction concepts, procedures, and their meaning. The exception to this was in directions to the student to retry the problem, which had minimal teacher input and thus limited mathematical connections. Lastly, when the teacher took over the problem solving via telling, it was less likely that their response contained mathematical connections.

Limitations

Before discussing the findings, I discuss a limitation to the present study stemming from the use of an instructional vignette in the study design. While vignettes and animated teaching simulations enable scholars to assess proxies for teacher practice in a standardized method and at scale, they preclude a contextualized and situated examination of the back-and-forth discourse that occurs between teachers and their students in classrooms. This limitation becomes most pronounced when the teacher employs a pedagogy of probing, for example, and the full course of dialogue is unavailable due to the vignette format. While one might conclude that a teacher who probes the student in a vignette study does not explicitly convey any mathematical connections in their response (e.g. “How did you come up with your strategy?”), the mathematical content could be known if the full dialogue with the student were possible. In this particular example of probing, I argue the teacher’s pedagogical approach preserves the cognitive demand of the task by holding the student accountable for high-level processes and products, despite the absence of explicit fractions connections. Thus, teacher maintenance of task cognitive demand must be understood as both multidimensional and circumscribed by the vignette format.

Additionally, while the use of an instructional vignette provides the benefit of standardizing the task and student misconception across all teachers, it precludes an analysis of the complex, situated nature of task implementation, such as a consideration of student sensemaking and sociomathematical norms in the classroom (Warshauer, 2015; Yackel & Cobb, 1996). This issue becomes apparent when examining the 17.5% of responses that were coded for an unspecified pedagogical approach and thus indeterminable level of student accountability. In these responses, it may very well be the case that teachers intended to enlist the student in problem solving, e.g. “Bring your work up to the board with me and let’s solve this problem together,” but the unspecified

nature of these approaches stemming from the vignette format which truncates the full course of dialogue precluded this analysis. Finally, the sample of teachers represented in this study leaned more male and more racially and ethnically diverse than the teacher workforce, limiting the generalizability of the findings.

Discussion

Overall, I found that nearly two-thirds of the responses coded for fractions content forged at least one connection to fraction concepts, procedures, and their underlying meaning. In most of these cases, the connection was established by relating the problem context to the fraction diagram. This finding could have simply been due to the diagram being embedded in the vignette itself. Still, given the prevalence of fraction diagrams for demonstrating part-whole representations in U.S. instruction and the importance of extending these representations to fraction operations (Siegler et al., 2010), this finding suggests the use of visual representations for fractions operations is low-hanging fruit for forging mathematical connections and thus preserving some of the mathematical integrity of tasks with high-level demands. Given that most fraction representations in elementary school center on part-whole visualizations, there is opportunity to expand on those representations to illustrate common fractions operations. For example, teachers can support students in making connections between procedures and their meaning by representing how to partition objects into smaller pieces for fraction division or redefine units for fraction multiplication (Siegler et al., 2010). These findings also show there is much room for teachers to forge connections between fraction operations and the measurement value of fractions, for example with the use of a number line to help struggling learners build greater fluency with fractions (Fuchs et al., 2013; Siegler et al., 2010).

Only a small portion of teachers made multiple connections among concepts, procedures, and meaning, and a substantial proportion did not make any connections at all, despite attending to

the mathematical content of the task. In instances where no connections were observed in teachers' treatment of fractions, this was typically due to an exclusive focus on the procedural steps in the standard algorithm for subtracting equivalent fractions with unlike denominators. These findings show that one mechanism by which the cognitive demands of high-level fractions tasks can devolve is through the reduction of mathematical connections. Prior work has shown a strict focus on procedural thinking is a common way students' cognitive processes decline while working on high-level math tasks (Henningsen & Stein, 1997). The present study builds on this work to show that teachers may similarly default to a strict focus on procedures in response to student misconceptions. Here, teachers may have felt the need to oversimplify the mathematics of the task to support the struggling student, inadvertently reducing the coherence of the mathematics being presented. As teachers may seek to reduce uncertainty and complexity in the classroom (e.g. Cohen, 2011), a strict focus on procedures may provide mathematical simplicity at the expense of rigor.

Teacher knowledge may also be at play, as prior work has demonstrated the links between teacher knowledge about common student struggles in fractions and their reported instructional responses (Copur-Gencturk et al., 2025). Future work in this area should further examine this domain-specific knowledge-action link with attention to mathematical connections. For example, through professional development focused on solving open-ended math tasks, elementary teachers can expand their notions about the nature of structure of mathematics and the variety of mathematical strategies used in the problem-solving process (Sevinc et al., 2025).

When it comes to the pedagogical strategies teachers used to hold the student accountable for the high-level processes or products of the task, I find that a majority of teachers took a teacher-directed problem-solving approach, resulting in a devolution of task cognitive demand. Prior research has also demonstrated the tendency for teachers to take over the problem solving in task-

based mathematics instruction (Middleton, 1999; Stein et al., 1996), suggesting the need for better teacher preparation in strategies like targeted probing that preserve problem-solving opportunities for students (e.g. Sullivan & Mornane, 2013).

I also find teachers rarely probed the student, aligning with prior work on U.S. teachers' use of probes when responding to students' mathematical errors or misconceptions. In a similar study that employed instructional vignettes, teachers rarely elicited or probed the student's mathematical thinking (Hill et al., in press). Prior research analyzing teachers' treatment of student errors has shown that Chinese teachers often took an inquiry-based approach, probing students about their errors to a greater extent as compared to U.S. teachers, which may help to explain the richer student discussion of mathematical concepts and higher mathematics achievement seen in the Chinese educational context (Schleppenbach et al., 2007). Teachers have also been found to provide probing guidance to students when they struggled with explaining and sensemaking; that is, teachers responded to different struggles with different pedagogical approaches (Warshauer, 2015). In the present study, the student struggle could be interpreted as conceptual, yet most teachers took a telling approach focused on steps to the standard algorithm. Future work on teacher maintenance of task cognitive demand might map out the relationship between teacher diagnoses for student fraction misconceptions and their ensuing instructional responses. Additionally, prior work shows the promise of video analysis for developing teacher notions of productive struggle (Warshauer et al., 2021), and the use of instructional vignettes such as the one featured in this study provides an option for a low-cost formative assessment tool enabling teacher educators to monitor the development of practices that preserve the demands of high-level tasks during implementation.

The present study also extends prior findings about the use of scaffolding to maintain the cognitive demand of high-level math tasks. Kokushkin and Tiraphatna (2020) define “productive

scaffolding” as a type of scaffold that preserves the cognitive demand of tasks, finding that the strategy of “quasi-induction” in the undergraduate setting enabled professors to teach mathematical proofs while maintaining task cognitive demand. In the present study, I find that the practice of scaffolding was more likely to occur in responses that drew connections between fraction concepts, procedures, and their meaning. Teachers’ productive use of scaffolds which enlist the student as a problem solver and orient their thinking towards mathematical connections may be a high-leverage practice that preserves task cognitive demand, enabling students to *do* mathematics in the face of challenges and misconception.

Notably, scaffolding and probing often co-occurred with privileging the student strategy, suggesting potential synergies between these pedagogical approaches. It could be the case that when noticing key components of the student strategy, teachers saw opportunities to leverage and build on student thinking via scaffolding, or to elicit and use student thinking further via probing. Indeed, prior work shows teachers are more responsive to a student’s struggle when using and implementing tasks that make student thinking—and the misconceptions therein—visible (Warshauer, 2015). More research is needed to understand the complex ways in which teacher noticing, content knowledge and knowledge about common student struggles in fractions, and student-centered instructional practice work together to formulate opportunities for students to *do* mathematics, with implications for teacher education and practice in the foundational domain of fractions.

Conclusion

Through this study, I aimed to understand the nature and quality of teacher treatment of fractions content and its connectedness and student accountability for problem solving when working on a fractions task with high-level demands. The findings illustrate the ways in which teachers can make connections between mathematical ideas and engineer student opportunities to

grapple with mathematics when confronted with a common student misconception about fractions. The present study highlights specific strategies teachers can use to cultivate student opportunities to learn fractions in K-8 mathematics, with implications for improving teacher knowledge, practice, and student learning in this crucial domain.

References

- Anderson, L. M. (1989). Classroom instruction. In M. C. Reynolds (Ed.), *Knowledge base for the beginning teacher* (pp. 101-115). Washington, DC: American Association of Colleges for Teacher Education.
- Behr, M. J., Post, T. R., Wachsmuth, I. (1986). Estimation and children's concept of rational number size. In Schoen, H. L. & Zweng, M. J. (eds.). *Estimation and mental computation: 1986 yearbook*. Reston, VA: The National Council of Teachers of Mathematics, Inc.
- Boston, M. (2012). Assessing Instructional Quality in Mathematics. *The Elementary School Journal*, 113(1), 76–104. <https://doi.org/10.1086/666387>
- Cazden, C. B. (1988). Classroom discourse: the language of teaching and learning. Portsmouth: Heinemann.
- Charalambous, C. Y., Agathangelou, S. A., Kasapi, E., & Christofidou, E. (2023). Learning to teach ambitiously: a multiple case study of practicing teachers' experimentation with enablers and extenders. *Journal of Mathematics Teacher Education*, 26(3), 363–394. <https://doi.org/10.1007/s10857-022-09532-9>
- Charalambous, C. Y. (2020). Reflecting on the troubling relationship between teacher knowledge and instructional quality and making a case for using an animated teaching simulation to disentangle this relationship. *ZDM Mathematics Education*, 52(2), 219-240. <http://dx.doi.org/10.1007/s11858-019-01089-x>
- Chazan, D., & Ball, D. (1999). Beyond being told not to tell. *For the Learning of Mathematics*, 19(2), 2–10.
- Cohen, D. K. (2011). *Teaching and its predicaments* (1st ed.). Harvard University Press.
- Copur-Gencturk, Y., Ezaki, J., & Jacobson, E. (2025). Missing link: How teachers' understanding of student common struggles is key to their instructional response. *Journal of Mathematics Teacher Education*. <https://doi.org/10.1007/s10857-025-09701-6>
- Copur-Gencturk, Y., & Doleck, T. (2021). Linking teachers' solution strategies to their performance on fraction word problems. *Teaching and Teacher Education*, 101, 103314. <https://doi.org/10.1016/j.tate.2021.103314>
- Copur-Gencturk, Y. (2021). Teachers' conceptual understanding of fraction operations: Results from a national sample of elementary school teachers. *Educational Studies in Mathematics*, 107(3), 525–545. <https://doi.org/10.1007/s10649-021-10033-4>
- Cramer, K., & Wyberg, T. (2009). Efficacy of Different Concrete Models for Teaching the Part-Whole Construct for Fractions. *Mathematical Thinking and Learning*, 11(4), 226–257. <https://doi.org/10.1080/10986060903246479>
- Deterding, N. M., & Waters, M. C. (2021). Flexible coding of in-depth interviews: A twenty-first-century approach. *Sociological Methods & Research*, 50(2), 708–739. <https://doi.org/10.1177/0049124118799377>

- Doyle, W. (1988). Work in mathematics classes: The context of students' thinking during instruction. *Educational Psychologist*, 23, 167-180.
- Estrella, S., Zakaryan, D., Olfos, R., & Espinoza, G. (2020). How teachers learn to maintain the cognitive demand of tasks through Lesson Study. *Journal of Mathematics Teacher Education*, 23(3), 293–310. <https://doi.org/10.1007/s10857-018-09423-y>
- Fazio, L., & Siegler, R. (2011). *Teaching Fractions. Educational Practices Series-22*. UNESCO International Bureau of Education.
- Fuchs, L. S., Schumacher, R. F., Long, J., Namkung, J., Hamlett, C. L., Cirino, P. T., Jordan, N. C., Siegler, R., Gersten, R., & Changas, P. (2013). Improving at-risk learners' understanding of fractions. *Journal of Educational Psychology*, 105(3), 683–700. <https://doi.org/10.1037/a0032446>
- Gearhart, M., Saxe, G. B., Seltzer, M., Schlackman, J., Ching, C. C., Nasir, N., Fall, R., Bennett, T., Rhine, S., & Sloan, T. F. (1999). Opportunities to Learn Fractions in Elementary Mathematics Classrooms. *Journal for Research in Mathematics Education*, 30(3), 286. <https://doi.org/10.2307/749837>
- Gresalfi, M., Martin, T., Hand, V., & Greeno, J. (2009). Constructing competence: An analysis of student participation in the activity systems of mathematics classrooms. *Educational Studies in Mathematics*, 70(1), 49–70. <https://doi.org/10.1007/s10649-008-9141-5>
- Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E., & Williamson, P. W. (2009). Teaching Practice: A Cross-Professional Perspective. *Teachers College Record: The Voice of Scholarship in Education*, 111(9), 2055–2100. <https://doi.org/10.1177/016146810911100905>
- Hansen, N., Jordan, N. C., & Rodrigues, J. (2017). Identifying learning difficulties with fractions: A longitudinal study of student growth from third through sixth grade. *Contemporary Educational Psychology*, 50, 45–59. <https://doi.org/10.1016/j.cedpsych.2015.11.002>
- Henningsen, M., & Stein, M. K. (1997). Mathematical tasks and student cognition: Classroom-based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for Research in Mathematics Education*, 28(5), 524–549.
- Herbst, P., Chazan, D., Chen, C. L., Chieu, V. M., & Weiss, M. (2011). Using comics-based representations of teaching, and technology, to bring practice to teacher education courses. *ZDM Mathematics Education*, 43, 91-103. <https://doi.org/10.1007/s11858-010-0290-5>
- Hiebert, J. & Carpenter, T. (1992). Learning and teaching with understanding. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 65–97). Macmillan.
- Hiebert, J. & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In J. Frank & K. Lester (Eds.), *Second handbook of research on mathematics teaching and learning* (pp. 371–404). Charlotte: Information Age.
- Hiebert, J. & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 1–27). Lawrence Erlbaum Associates, Inc.
- Hill, H. C. Garcia Coppersmith, J., Kleen, H., Pollard, C. (in press). Measuring Standards-Based Instructional Practices at Scale: A New Approach. *Educational Assessment*.

- Kokushkin, V., & Tiraphatna, M. (2020). An Instructor's Actions for Maintaining the Cognitive Demands of Tasks in Teaching Mathematical Induction. *Conference Papers -- Psychology of Mathematics & Education of North America*, 2092–2096.
<https://doi.org/10.51272/pmena.42.2020-352>
- Ma, L. (2006). "Knowing and teaching fractions: A cross-cultural study of American and Chinese mathematics teachers." *Contemporary Educational Psychology*, 31(4), 438–457.
- Mack, N. K. (2001). Building on informal knowledge through instruction in a complex content domain: partitioning, units, and understanding multiplication of fractions. *Journal for Research in Mathematics Education*, 32(3), 267–296.
- Mack, N. K. (1993). Learning rational numbers with understanding: The case of informal knowledge. In T. P. Carpenter, E. Fennema, & T. A. Romberg (Eds.), *Rational numbers: An integration of research* (pp. 85–105). Hillsdale, NJ: Lawrence Erlbaum Associate
- Middleton, J. A. (1999). Curricular influences on the motivational beliefs and practice of two middle school mathematics teachers: A follow-up study. *Journal for Research in Mathematics Education*, 30(3), 349–358
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2020). *Qualitative data analysis: A methods sourcebook* (4th ed.). SAGE.
- Murphy, C., Hatisaru, V., & Chick, H. (2024). Editorial for a special issue on mathematical connections in the teaching and learning of mathematics. *Research in Mathematics Education*, 26(2), 231–236. <https://doi.org/10.1080/14794802.2024.2372378>
- National Center for Education Statistics [NCES]. (2023). *Characteristics of public school teachers*. Condition of education. U.S. Department of Education, Institute of Education Sciences.
<https://nces.ed.gov/programs/coe/indicator/clr>
- National Center for Education Statistics [NCES]. (2003). *Teaching mathematics in seven countries: results from the TIMSS video study*. Washington, DC: U.S. Department of Education.
- National Council of Teachers of Mathematics. (2007). *The learning of mathematics: 69th NCTM yearbook*. Reston, VA: National Council of Teachers of Mathematics.
- National Governors Association Center for Best Practices & Council of Chief State School Officers (2010). *Common Core State Standards for Mathematics*. Washington, DC.
- Newton, K. J. (2008). An Extensive Analysis of Preservice Elementary Teachers' Knowledge of Fractions. *American Educational Research Journal*, 45(4), 1080–1110.
<https://doi.org/10.3102/0002831208320851>
- O'Connor, M. C., & Michaels, S. (1993). Aligning academic task and participation status through revoicing: Analysis of a classroom discourse strategy. *Anthropology and Education Quarterly*, 24(4), 318–335.
- Olanoff, D., Lo, J.-J., & Tobias, J. (2014). Mathematical Content Knowledge for Teaching Elementary Mathematics: A Focus on Fractions. *The Mathematics Enthusiast*, 11(2), 267–310.
<https://doi.org/10.54870/1551-3440.1304>

- Rittle-Johnson, B., & Schneider, M. (2015). Developing conceptual and procedural knowledge of mathematics. In R. C. Kadosh & A. Dowker (Eds.), *The Oxford handbook of numerical cognition* (pp. 1118–1134). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199642342.013.014>
- Rittle-Johnson, B., Siegler, R. S., Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of educational psychology*, 93, 346–362.
- Schleppenbach, M., Flevares, L. M., Sims, L. M., & Perry, M. (2007). Teachers’ Responses to Student Mistakes in Chinese and U.S. Mathematics Classrooms. *The Elementary School Journal*, 108(2), 131–147. <https://doi.org/10.1086/525551>
- Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sensemaking in mathematics. In D. Grouws (Ed.). New York: MacMillan
- Schoenfeld, A., Fink, H., Zuñiga-Ruiz, S., Huang, S., Wei, X., & Chirinda, B. (2023). *Helping students become powerful mathematics thinkers: Case studies of teaching for robust understanding*. <https://doi.org/10.4324/9781003375197>
- Sevinc, S., Cross Francis, D., Hudson, R., & Liu, J. (2025). The development of elementary teachers’ personal meanings of problem-solving through engaging in open-ended tasks. *Mathematics Education Research Journal*, 37(2), 341–367. <https://doi.org/10.1007/s13394-024-00493-1>
- Shaughnessy, M., & Boerst, T. A. (2018). Uncovering the skills that preservice teachers bring to teacher education: The practice of eliciting a student’s thinking. *Journal of Teacher Education*, 69(1), 40-55.
- Siegler, R., Carpenter, T., Fennell, F., Geary, D., Lewis, J., Okamoto, Y., Thompson, L., & Wray, J. (2010). *Developing Effective Fractions Instruction for Kindergarten through 8th Grade. IES Practice Guide. NCEE 2010-4039*. What Works Clearinghouse. <https://eric.ed.gov/?id=ED512043>
- Silver, E. A., & Stein, M. K. (1996). The Quasar Project: The “Revolution of the Possible” in Mathematics Instructional Reform in Urban Middle Schools. *Urban Education*, 30(4), 476–521. <https://doi.org/10.1177/0042085996030004006>
- Smith, J. P., III. (2002). *The Development of Students’ Knowledge of Fractions and Ratios*. National Council of Teachers of Mathematics.
- Star, J. R. (2005). Reconceptualizing procedural knowledge. *Journal for Research in Mathematics Education*, 36, 404–411.
- Stein, M. K., Grover, B. W., & Henningsen, M. (1996). Building Student Capacity for Mathematical Thinking and Reasoning: An Analysis of Mathematical Tasks Used in Reform Classrooms. *American Educational Research Journal*, 33(2), 455-488. <https://doi-org.ezp-prod1.hul.harvard.edu/10.3102/00028312033002455>
- Stein, M. K., & Lane, S. (1996). Instructional tasks and the development of student capacity to think and reason: An analysis of the relationship between teaching and learning in a reform mathematics project. *Educational Research and Evaluation*, 2, 50 – 80.
- Stein, M. K. & Smith, M. S. (1998). “Mathematical tasks as a framework for reflection: From research to practice.” *Mathematics teaching in the middle school* 3.4: 268-275.

- Stein, M. K., Smith, M. S., Henningsen, M. A., & Silver, E. A. (2009). *Implementing standards-based math instruction: A casebook for professional development*. Teachers College Press.
- Stewart, E. & Ball, L. (2023). The Tension Between Allowing Student Struggle and Providing Support When Teaching Problem-Solving in Primary School Mathematics. *Canadian Journal of Science, Mathematics, and Technology Education*, 23, 791–817. <https://doi.org/10.1007/s42330-024-00309-1>
- Stigler, J. W., & Hiebert, J. (2004). Improving mathematics teaching. *Educational Leadership*, 61(15), 12-16.
- Sullivan, P., & Mornane, A. (2013). Exploring teachers' use of, and students' reactions to, challenging mathematics tasks. *Mathematics Education Research Journal*, 25, 1-21.
- Sullivan, P., Mousley, J., & Zevenbergen, R. (2006). Teacher Actions to Maximize Mathematics Learning Opportunities in Heterogeneous Classrooms. *International Journal of Science and Mathematics Education*, 4(1), 117–143. <https://doi.org/10.1007/s10763-005-9002-y>
- Terwel, J., Van Oers, B., Van Dijk, I., & Van Den Eeden, P. (2009). Are representations to be provided or generated in primary mathematics education? Effects on transfer. *Educational Research and Evaluation*, 15(1), 25–44. <https://doi.org/10.1080/13803610802481265>
- Warshauer, H. K., Starkey, C., Herrera, C. A., & Smith, S. (2021). Developing prospective teachers' noticing and notions of productive struggle with video analysis in a mathematics content course. *Journal of Mathematics Teacher Education*, 24(1), 89–121. <https://doi.org/10.1007/s10857-019-09451-2>
- Warshauer, H. K. (2015). Productive struggle in middle school mathematics classrooms. *Journal of Mathematics Teacher Education*, 18(4), 375–400. <https://doi.org/10.1007/s10857-014-9286-3>
- Yackel, E., & Cobb, P. (1996). Sociomathematical Norms, Argumentation, and Autonomy in Mathematics. *Journal for Research in Mathematics Education*, 27(4), 458. <https://doi.org/10.2307/749877>

Appendix A3

Table A3.1. Kappa Scores for Fractions Content Codes

	Fractions Content Code	Kappa	Positive Agreement
<i>Connects Fraction Concepts, Procedures, and/or Their Underlying Meaning</i>	Standard algorithm for creating equivalent fractions focused on the meaning behind the steps	0.75	0.78
	Understanding the numerator and as values in relation	0.65	0.68
	Connects the visual representation to the problem context	0.79	0.85
	Estimating/ intuiting the value of $\frac{7}{8}$, i.e. that it is nearly one whole	0.90	0.90
<i>Does Not Connect Fraction Concepts, Procedures, or Their Underlying Meaning</i>	Standard algorithm for creating equivalent fractions focused on operations	0.79	0.84
	The fact that $\frac{7}{8}$ is not the same as $\frac{7}{16}$ without elaboration	0.82	0.86
<i>Other (fractions content and connections undetermined)</i>	No math content	0.75	0.78
	Mathematical errors and/or imprecision	0.71	0.77
	Unspecified mathematical focus	0.78	0.84

Note: $N=228$ responses from $N=228$ teachers. Positive agreement describes the two raters' agreement on the subset of responses excluding those in which neither rater applied the code.

Table A3.2. Kappa Scores for Codes for Student Accountability

Student accountability for high-level process or product	Pedagogy Code	Kappa	Positive Agreement
<i>High</i>	Probing	1.00	1.00
<i>High</i>	Direction to retry	0.69	0.73
<i>Moderate</i>	Scaffolding	0.67	0.77
<i>Moderate</i>	Privileging student strategy	0.82	0.86
<i>Low</i>	Telling	0.67	0.85
<i>N/A</i>	Unspecified pedagogical approach	0.85	0.88

Note: $N = 228$ responses from $N = 228$ teachers. Positive agreement describes the two raters' agreement on the subset of responses excluding those in which both raters did not apply the code

Conclusion

The findings of this dissertation project contribute to a better understanding of mechanisms behind racialized achievement disparities and underperformance in the (post-)reform era of K-8 mathematics teaching and learning. The first study examined teacher expectations, judgments, and responsiveness to students' correct and non-standard mathematical ideas. The second and third study focused on teacher explanations and responses to common student misconceptions during the critical implementation stage of task-based mathematics instruction. All studies leveraged approximations of practice via instructional vignettes to capture action-related competencies of core dimensions of ambitious mathematics teaching (Charalambous, 2020).

Summary of Findings

Findings from the first study show teachers expected greater task difficulty in both the Latinx/e and Black classroom conditions relative to the White, and no significant differences by condition in dimensions of responsive teaching. However, we found trend-level evidence to suggest teachers may be more likely to support student sense-making and provide more positive, substantive affirmations to Black students relative to White students for the same mathematical solution.

The second study similarly showed teachers to be relatively equitable in their causal attributions for and instructional responses to student struggles during the implementation of high-level mathematics tasks. Moreover, I found teacher endorsement of statements attributing the student struggle to internal factors—but not external factors—positively related to their preferences to alter task cognitive demand. Looking at teacher and school context factors, I find those who are women and those teaching more advanced grade levels were less likely to endorse internal causal attributions for the student struggle. Lastly, teachers' instructional preferences to routinize,

remediate, and extend a math task did not differ by student race-ethnicity but did differ descriptively according to the level of racial-ethnic segregation in the teaching context.

The final study examined teacher treatment of mathematics content and student accountability for high-level processes and products when responding to a hypothetical student's misconception on a high-level fractions task. I found that most teachers took a teacher-directed approach, relying on standard algorithmic steps rather than connections between mathematical ideas. Still, many teachers forged at least one mathematical connection, and it was most often through leveraging a visual fractions representation. Pedagogies of scaffolding and probing often co-occurred with privileging the student strategy, suggesting a potentially complimentary role between these pedagogical approaches.

Discussion

Fundamentally, this set of studies is concerned with aspects of ambitious mathematics teaching that center the student as an agentic sense maker and enlist them in the doing of mathematics. Taken together, these three dissertation papers depict a nuanced picture of biases and practice in reform-oriented K-8 mathematics teaching. Teachers in this study did not, on average, exhibit racial-ethnic biases in their evaluation of students' mathematical work, their use of student thinking, their causal attributions for a struggling student, or in their reported preferences for altering the cognitive demand of the task. Yet their expectations about the difficulty of that same work for "typical" students were racialized in ways consistent with longstanding stereotypes about Black and Latinx/e learners, and their instructional responses revealed persistent tendencies to diminish the cognitive demand of a high-level fractions task by winnowing down mathematical connections and assuming control of problem solving over the student.

Collectively, the findings suggest a potential instructional mechanism behind racialized achievement disparities in U.S. mathematics may be task selection, informed by biased expectations, rather than teacher responsiveness or task modification. Theoretically, the findings suggest that biases may be most likely to surface in generalized expectations and in contexts where race is made salient alongside sociocultural markers, and that individuating information about student competence can partially interrupt these processes. Practically, they point to the need for teacher learning experiences that (a) surface and disrupt racialized expectations about mathematical difficulty; (b) re-orient causal attributions toward opportunity-to-learn frameworks that position instruction as a powerful lever; and (c) strengthen teachers' capacity to maintain the cognitive demand of high-level tasks by building on student thinking, forging rich mathematical connections—especially in foundational domains like fractions—and holding students accountable for doing mathematics.

Limitations

A limitation of this dissertation stems from the use of instructional vignettes that approximate practice. In using an experimental design to examine the main effects of student race-ethnicity, this set of studies may have sacrificed external and ecological validity in its effort to maintain internal validity. Teaching is fundamentally socially situated, and the vignette study format limits an examination of ambitious teaching in context. Future studies may employ new methodologies or adapt existing ones to use experimental designs while situating teaching in context. Moreover, emerging natural language processing and machine learning techniques that enable scholars to process and examine large datasets, including classroom transcripts, may better capture the complex dynamics of classroom practices, student thinking, and how they relate to other factors to support mathematics learning over time.

Implications

The findings of this dissertation project have implications for research and teacher training in ambitious mathematics teaching. The first study demonstrates the need for anti-bias training to counter stereotypes about racially minoritized learners and the difficulties they are expected to encounter on mathematics problems. In the following study, we see that student-facing teacher explanations for mathematical challenges relate to teacher practice more than external-facing explanations. This suggests factors like student practice, readiness, and engagement in mathematics inform instructional action over external factors outside of teachers' control. As such, teacher education and development focused on diagnosing student challenges in terms of an internal locus may be fruitful for designing instructional interventions in task-based mathematics learning.

Finally, as challenging and complex as task-based mathematics instruction may be, especially in the face of student misconceptions and impasses, key practices show promise for preserving the cognitive demand of high-level tasks. Teacher practices that forge connections between mathematical concepts, procedures, and their underlying meaning—such as with visual diagrams to imbue fractions procedures with meaning—can preserve the high-level demands of tasks when resolving student misunderstandings. Teacher education and training focused on making connections as well scaffolding, probing, and privileging student strategies may support educators to implement high-level tasks in a way that preserves mathematical coherence and student accountability for doing mathematics.

Overall, this dissertation project contributes to a deeper understanding of teacher judgments and decision making in key dimensions of ambitious mathematics teaching. The findings build on a vision for mathematics learning experiences that challenge and engage all learners.

References

- Boston, M. D., & Wilhelm, A. G. (2017). Middle School Mathematics Instruction in Instructionally Focused Urban Districts. *Urban Education*, 52(7), 829–861. <https://doi.org/10.1177/0042085915574528>
- Charalambous, C. Y. (2020). Reflecting on the troubling relationship between teacher knowledge and instructional quality and making a case for using an animated teaching simulation to disentangle this relationship. *ZDM Mathematics Education*, 52(2), 219-240. <http://dx.doi.org/10.1007/s11858-019-01089-x>
- Lampert, M., Franke, M. L., Kazemi, E., Ghouseini, H., Turrou, A. C., Beasley, H., Cunard, A., & Crowe, K. (2013). Keeping it complex: Using rehearsals to support novice teacher learning of ambitious teaching. *Journal of Teacher Education*, 64(3), 226–243. <https://doi.org/10.1177/0022487112473837>
- National Council of Teachers of Mathematics (NCTM). (2000). *Principles and standards for school mathematics*. NCTM.
- National Council of Teachers of Mathematics (NCTM). (2014). *Principles to actions: ensuring mathematical success for all*. NCTM.
- National Governors Association Center for Best Practices & Council of Chief State School Officers (2010). *Common Core State Standards for Mathematics*. Washington, DC
- Smylie, M., & Wenzel, S. (2006). *Promoting instructional improvement: A strategic human resource management perspective*. Consortium on Chicago School Research.
- The New Teacher Project (TNTP) (2018). *The opportunity myth*. New York, NY. Retrieved September 20, 2018, from https://tntp.org/assets/documents/TNTP_The-Opportunity-Myth_Web.pdf
- Wilhelm, A. G., Munter, C., & Jackson, K. (2017). Examining Relations between Teachers' Explanations of Sources of Students' Difficulty in Mathematics and Students' Opportunities to Learn. *The Elementary School Journal*, 117(3), 345–370. <https://doi.org/10.1086/690113>